

The Reduced Partitioning Procedure in
Configuration Interaction Studies

By

RODNEY JOSEPH BARTLETT

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1971

DEDICATED

to

MY PARENTS

ACKNOWLEDGEMENTS

It is my pleasure to gratefully acknowledge my chairman, Prof. Yngve Öhrn, for the unfaltering assistance, encouragement, and supervision he has given me during the period I have spent in the Quantum Theory Project. Without his cooperation in all manner of things, my stay here would have been much less rewarding. Also, I wish to sincerely express my appreciation to Prof. Per-Olov Löwdin, my co-chairman, for his intellectual stimulation and for several conversations in which he made numerous fruitful comments which improved the quality of this dissertation. To Dr. Erkki Brändas, I owe a special word of thanks, since this research was undertaken with his assistance and would have been impossible without his guidance. To say that the exchange of ideas between us was of great benefit to me is an understatement. I also wish to express my appreciation to the other members of my supervisory committee for their continued cooperation.

All of the members of the Quantum Theory Project have been helpful to me, but particularly I would like to thank Dr. C. E. Reid and Dr. Joe Kouba for assistance in the use of some of their computer programs, as well as Nelson Beebe and John Bellum for additional computational aid. I am also indebted to Dr. Micheal Hayns for his careful drawing of the figures.

Lastly, I am especially grateful to Beverly Bartlett for her tireless efforts in the typing of this dissertation as well as her constant encouragement during its preparation.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii
Chapter	
I. INTRODUCTION	1
II. THE PARTITIONING TECHNIQUE AND PERTURBATION THEORY	10
II.1 The Partitioning Technique	10
II.2 The Partitioning Technique in Operator Form	13
II.3 Brillouin-Wigner Perturbation Theory	15
II.4 Rayleigh-Schrödinger Perturbation Theory	18
II.5 Inner and Outer Projections	20
III. PÂDÉ APPROXIMANTS	24
III.1 Deviation and Definitions	24
III.2 Connection with Moment Expansions and Inner Projections	29
IV. THE REDUCED PARTITIONING PROCEDURE	31
IV.1 Non-Linear Summation of Perturbation Expansions	31
IV.2 A Model Hamiltonian for a Finite Basis	35
IV.3 The Fundamental Equations for the Reduction Process	39
IV.4 The Method of Steepest Descent and the Geometric Sumrule	45

V.	EXCITED STATES	50
V.1	The Super-Matrix Formulation	50
V.2	The Selective Reduction Procedure	56
V.3	The Geometric Formula for a Multi-Dimensional Reference Space	59
VI.	NUMERICAL APPLICATIONS OF THE REDUCED PARTITIONING PROCEDURE	62
VI.1	Summary of the Computational Procedure	63
VI.2	Results and Discussion: The Single Partitioning Case	69
VI.3	Results and Discussion: The Selective Reduction Procedure for Multiple States	87
	BIBLIOGRAPHY	106
	BIOGRAPHICAL SKETCH	110

LIST OF TABLES

Table	Page
1. Reduction Energies for H_2 with the Coulson Reference Function	71
2. Reduction Energies for H_2 with a Hartree-Fock Reference Function	73
3. Reduction Energies for HeH^+ with a Hartree-Fock Reference Function	75
4. Reduction Energies for the $X^1\Sigma^+$ State of HeH^+ as a Function of Internuclear Separation	77
5. Spectroscopic Parameters for Different Approximations to the $X^1\Sigma^+$ State of HeH^+	85
6. Illustration of Selective Reduction for Multiple States of HeH^+	88
7. Reduction Energies for the $A^1\Sigma^+$ State of HeH^+ as a Function of Internuclear Separation	91
8. Comparison of Rigorous and Approximate Multiple State Reduction Energies for HeH^+ as a Function of Internuclear Separation	97
9. Reduction Energies for the $^1\Sigma^+$ Excited States of HeH^+ as a Function of Internuclear Separation	101

LIST OF FIGURES

Figure	Page
1. A Comparison of the HF and Full CI Potential Curves for the Ground ($X^1\Sigma^+$) State of HeH^+	81
2. Low Order Reduction Energies for the $X^1\Sigma^+$ State of HeH^+	83
3. Full CI Results for the Ground and First Excited States of HeH^+	93
4. Low Order Reduction Energies for the $A^1\Sigma^+$ State of HeH^+	95
5. Improved Excited $^1\Sigma^+$ States for HeH^+	100
6. Reduction Energies for the $B^1\Sigma^+$ State of HeH^+	104

Abstract of Dissertation Presented to the
Graduate Council of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

THE REDUCED PARTITIONING PROCEDURE IN
CONFIGURATION INTERACTION STUDIES

By

Rodney Joseph Bartlett

June, 1971

Chairman: Prof. Yngve Öhrn

Co-Chairman: Prof. Per-Olov Löwdin

Major Department: Chemistry

Within the configuration interaction (CI) framework, a higher order perturbation approach has been developed which allows a drastic reduction in the number of n-particle functions required to give the energy of a CI wavefunction to a desired numerical accuracy. By means of this method, significant improvements over the standard Hartree-Fock (HF) and truncated CI calculations are obtainable. Since the procedure does not require solutions of large secular equations, the technique also has implications for full CI calculations. As this procedure complements the natural orbital methods, the two techniques can be used in conjunction for additional flexibility.

The theory derived to effect the reduction process, i.e., the reduced partitioning procedure, employs the techniques of inner and outer projections and Padé approximants to yield a tractable computational scheme with known convergence properties. The theory is formulated both for a single unperturbed reference function and for a multi-dimensional reference space to use in applications to excited states. In the latter case, a method

which uniformly treats the excited states as well as one which permits a selective reduction for one specific state is obtained. These reduction equations are also shown to be related to the method of steepest descent.

The first order solution is of particular significance. In the examples studied, this result, which corresponds to the application of a geometric sumrule to the perturbation expansion, is found to give sixty to seventy per cent of the possible energy improvement obtainable from the specified basis. This solution also has some computationally simplifying features when the reference function is of the HF or CI variety.

Preliminary applications to the H_2 molecule and the HeH^+ molecular ion tend to justify the reduction procedure for more general cases. In these examples, no more than a twelfth order solution is required to give essential agreement with the full CI result in each case, even though as many as eighty-two configurations constitute the full n-particle basis. In addition, it is found that only a third order solution is required to account for more than ninety per cent of the possible correlation improvement over a HF calculation. These calculations presented here also include potential curves for the HF, the full CI, and the perturbation corrections for selected states of HeH^+ .

CHAPTER I. INTRODUCTION

The majority of theoretical studies of atoms, molecules, and solids has its origin in the time dependent, non-relativistic Schrödinger equation [1],

$$H\psi = i\hbar \frac{\partial \psi}{\partial t} . \quad (1.1)$$

More commonly, though, one focuses on its progeny, the time-independent form

$$H\Psi_n = E_n \Psi_n , \quad (1.2)$$

which defines the "stationary state" eigenfunctions, $\{\Psi_n\}$. The complete Hamiltonian operator for a given problem is an enormously complex entity, whose eigenfunctions would contain all electronic, vibrational, and rotational motions that are possible for the system as well as a variety of hyper-fine interactions. This enigma is never considered in its entirety, but, subject to the problem of interest, one makes certain assumptions as to the form of H . Even then the equations obtained are probably still partially coupled and require another approximation to effect a separation. It is this process and the Born-Oppenheimer, or "clamped nuclei," approximation [2] that lead us to the usual Hamiltonian for the electronic motion,

$$H = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{i < j} \frac{1}{r_{ij}} - \sum_{\alpha} \sum_i z_{\alpha} / r_{\alpha i} + \sum_{\alpha < \beta} z_{\alpha} z_{\beta} / R_{\alpha \beta} . \quad (1.3)$$

The quantities in (1.3) are the kinetic energy operator and terms arising from the coulombic potentials for electron-electron, electron-nuclear, and

nuclear-nuclear interactions. Within the clamped nucleii approximation the last term is merely a constant. When additional interactions are to be considered, they are customarily treated as a perturbation of (1.3).

Even though (1.3) is a distinct simplification, it is still impossible to solve (1.2) for a system with more than one electron, necessitating approximate solutions as the only recourse. In this respect, powerful methods have been developed to approximately solve (1.2). These fall primarily into two categories, variational methods, based on the Ritz variational principle [3], and perturbation theory. The former has traditionally been the more important of the two, with the latter essentially limited to the effects of external fields. However, a renewed interest in perturbation methods is very much in evidence currently.

Since (1.3) is bounded from below, the variational theorem allows one to approximate Ψ by some $\bar{\Psi}$ which contains several parameters to be optimized, form the functional

$$\bar{E} = \langle \bar{\Psi} | H | \bar{\Psi} \rangle / \langle \bar{\Psi} | \bar{\Psi} \rangle , \quad (1.4)$$

and by minimizing \bar{E} with respect to the parameters embedded in $\bar{\Psi}$, determine an approximate wavefunction and an $\bar{E} \geq E$. With a sufficiently flexible expression for $\bar{\Psi}$, one can thusly obtain as good a solution as desired. In this manner the energies of a few simple two- and three-electron systems have been calculated to experimental or better accuracy.

For less simple systems, two relatively general variational methods have replaced the "special case" solutions of the previous period. These are the Hartree-Fock method [4], which for molecules is invariably implemented as formulated by Roothaan [5], and the method of superposition of configurations or configuration interaction. The ordinary Hartree-Fock (HF) wavefunction for an atom or molecule with n electrons is assumed to

be a single Slater determinant,

$$\Psi_{HF} = \langle \psi_1(1) \psi_2(2) \dots \psi_n(n) \rangle \quad (1.5)$$

composed of n one-particle functions (spin-orbitals) where each spin-orbital is a spatial orbital ξ_i associated with either α or β spin. In the Roothaan procedure, the functions $\{\xi_i\}$ are assumed to be approximated by expanding in a given set of functions $\{x_\mu\}$,*

$$\xi_i = \sum_\mu x_\mu c_{\mu i} \quad (1.6)$$

with arbitrary expansion coefficients $c_{\mu i}$. The coefficients can be determined by the variational principle (1.4) to give the lowest possible energy for the single determinant form. Besides being conceptually satisfying, a wavefunction of this form possesses a degree of justification for the prediction of ionization potentials and one-electron properties due to the Koopmans [7] and Moeller-Plesset [8] theorems. On the other hand, all single determinant wavefunctions are unable to properly treat the instantaneous electron interactions, i.e., "correlation," which arises from the second term of (1.3). This has several consequences, one of which is that a standard Hartree-Fock wavefunction for a molecule cannot generally be used to obtain reasonable potential energy curves since it will not dissociate properly except in a few very special cases. Furthermore, one encounters difficulties with symmetry in proceeding to open shell studies resulting in the near impossibility of treating excited states.

The configuration interaction (CI) procedure effectively eliminates the complications inherent in the Hartree-Fock scheme, but not without

*A common basis set is the Slater type orbitals [6], $x(l,m,n) = r^{n-1} e^{-\alpha r} Y(l,m)$, where l, m , and n are the usual quantum numbers, r the magnitude of the radial coordinate relative to an arbitrary origin, $Y(l,m)$ the spherical harmonics, and α the orbital exponent which is to be selected in some manner.

causing some new problems of its own. If one considers the expression (1.6) and if there are m expansion functions $\{x_\mu\}$, then with spin, it follows that $2m(\psi_i)$ are possible. Consequently, if $2m > n$, many more than one determinant may be generated from these $\{\psi_i\}$ by forming the $\binom{2m}{n}$ combinations. In the CI method one can consider these determinants, $\{D_k\}$, to be the basis for a trial wavefunction,

$$\tilde{\Psi} = \sum_k D_k C_k , \quad (1.7)$$

with the $\{C_k\}$ to be determined from (1.4). Other than the multi-determinantal nature of the CI approach, which permits electron correlation, the primary distinction between it and Hartree-Fock theory is that the expansion coefficients in the $\{\xi_i\}$ are not usually determined via an energy criterion but rather fixed -- such as by orthogonalization of the $\{x_\mu\}$ -- thereby placing the flexibility of the wavefunction exclusively within the $\{C_k\}$. (Excluding the multi-configurational Hartree-Fock techniques, the one exception to this statement involves natural orbital iterations as discussed below.)

Immediately, a problem with the expression (1.7) becomes apparent, since the number of determinants $\binom{2m}{n}$ rapidly becomes astronomical even for a relatively modest basis set, $\{x_\mu\}$. If one is interested in a state or states of a certain symmetry, many of these determinants can be excluded by group theory since they will contain no component of that state. Even better, the individual determinants which do contribute to the state of interest can be grouped into linear combinations, $\{B_L\}$, such that each B_L is a pure symmetry function. By so doing, the expansion (1.7) can be substantially shortened, but, even so, the number of configurations is still likely to be very large.

If one forms the approximate wavefunction

$$\tilde{\Psi} = \sum_L B_L C_L , \quad (1.8)$$

by including every configuration that can be constructed from the given basis $\{x_\mu\}$, then the result obtained from the variational principle is termed the "full" CI solution for the problem. This solution has some convenient properties. First, it is invariant to any transformation among the basis functions $\{x_\mu\}$ and thus is independent of the explicit form for the $\{\xi_i\}$; or, equivalently, the final result is solely determined by the initial basis $\{x_\mu\}$. The second advantage is that by using all possible $\{B_L\}$, no dubious selection of configurations needs to be made. Unfortunately, though, the number of configurations needed for the full CI is still usually much too large to enable one to use all of them, a fact which leads to severe diagonalization problems among others. In order to avoid most of these problems, it is necessary to make a somewhat arbitrary selection of the configurations thought to be important and then solve this "truncated" CI problem. If so, the full CI features are no longer generally true. In particular, the explicit combinations $\{\xi_i\}$ are important to the results obtained.

A second problem encountered, the choice of basis $\{x_\mu\}$, is a universal one for all methods with the exception of purely numerical solutions. In some ways, however, the selection is more crucial for CI studies than in others, such as the HF. The reason for this is two-fold. First, contrary to a HF calculation, if the CI procedure is employed, one is generally interested in excited states as well as the ground state. This usually requires a basis that can do both adequately, a task which is exceedingly difficult to manage. Also, there is little or no "chemical intuition" which can suggest orbital exponents that will allow a satisfactory description of excited states. Of course, if the basis set were large enough, then the choice of the $\{x_\mu\}$ would not be as crucial, and this is a fairly

reasonable alternative in HF theory, but in CI one is again led to the problem with the number of possible configurations.

One has essentially attempted to resolve this dilemma with two methods, both used separately and in conjunction. These are orbital exponent optimization, which is most effective when undertaken for each state independently, and natural orbital iterations. The first of these is self-explanatory, although still largely prohibitive due to the extreme cost. The second is a potent tool for improving less than full CI calculations by exploiting the flexibility in the selections of the linear combinations $\{\xi_i\}$. This technique makes use of the concept of "natural spin orbitals," introduced by Löwdin [9], which are defined as the eigenvectors of the first order reduced density matrix,

$$\gamma(x_1 | x'_1) = \int \bar{\Psi}^*(x_1 x_2 \dots x_n) \bar{\Psi}(x'_1 x_2 \dots x_n) dx_2 \dots dx_n \quad (1.9)$$

in the sense of

$$\int \gamma(x_1 | x'_1) \varphi_k(x'_1) dx'_1 = n_k \varphi_k(x_1) , \quad (1.10)$$

where n_k is the occupation number associated with the natural spin orbital φ_k . From a theorem due to Schmidt and later Coleman [10], it is found that this set of orbitals has a certain optimum property. That is, for a full CI wavefunction, the selection of the natural orbitals with largest occupation numbers as the spin orbitals from which the configurations are constructed results in a much shorter approximate expansion which has the property of being in maximum possible coincidence with the full CI [11]. Natural orbital iterations further utilize this feature, except now pertaining to a truncated CI. Since such an approximation is not invariant to the choice of the $\{\xi_i\}$, by obtaining a set of approximate natural

orbitals from the density matrix associated with this CI wavefunction, one can select a set via an occupation number criterion from which another trial CI wavefunction is generated. This set is expected to be an improvement over the previous $\{\xi_i\}$. Consequently, by continuing the process and assuming convergence, the mixing of the basis functions $\{x_\mu\}$ is thusly "optimized" [12,13]. This technique is very useful, but it also has deficiencies, one of which is the dependence on the initial selection of configurations.

Having outlined the essential components of the two most important quantum chemical methods, given a satisfactory basis set, the predominant problem that emerges in progressing beyond the Hartree-Fock approximation, as well as utilizing configuration interaction to maximum efficiency, lies in the exorbitant numbers of configurations that need to be considered. The treatment of this standard problem in various contexts is the objective of this investigation. By appealing to the partitioning technique [14,15] and perturbation theory, we shall develop an approach to this problem that differs from the natural orbital methods, although it is entirely complementary. This process, the "reduced partitioning procedure," shows promise of being useful in a variety of applications within the configuration interaction domain.

The basic idea of the procedure is to reduce the dimension of a CI problem by grouping N configurations into M functions, such that the energy can be obtained to a desired numerical accuracy even though $M \ll N$. The main distinction between this "reduction" procedure and the natural orbital methods is that the M functions obtained are linear combinations of the previous configurations, whereas a natural orbital "reduction" replaces one set of configurations with a second. Thus, since each technique emphasizes different aspects of the problem, they can be used in conjunction for additional flexibility.

The method employed to achieve this result is essentially perturbation theory, but in a more general framework than is customary. Traditionally, the feasibility of applying perturbation theory to problems has depended on three elements: the existence of an exact solution to an unperturbed Hamiltonian; the convergence of a linear perturbation expansion; and, due to inherent computational problems, the suitability of first and second order solutions. These are impractical restrictions for arbitrary, many-electron systems, however, and preferably variational principles should be incorporated with perturbation techniques to give more powerful methods for the calculation of molecular properties and energies. The perturbation approach we have formulated includes the possibility of using any approximate wavefunction as the unperturbed solution, guarantees convergence via non-linear summations, and allows higher order solutions to be easily obtained.

The reduced partitioning procedure (RPP) will have several advantages over the more standard methods. For instance, a full CI calculation is seldom soluble since the diagonalization of large Hamiltonian matrices, extremely time consuming at best, is usually impossible due to near linear dependencies. Although the reduction process would generally require that all the matrix elements are constructed, it would only require a diagonalization of an $M \times M$ Hamiltonian matrix for $M \ll N$, thereby providing a feasible alternative to the full CI solution. Even though a full CI may not be the objective, one may be interested in certain intermediate solutions which would give a substantial improvement over a Hartree-Fock or truncated CI wavefunction. In this respect, also, the RPP suggests some possibilities. One of these, the first order solution, yields a dramatic energy decrease while possessing important computational simplifications. This first order solution can also be connected with the

"geometric" approximation and the method of "steepest descent." For excited states, the RPP will provide two alternative approaches which have different advantages. One of these allows a selective treatment of the particular state of interest. The question of properties other than the energy is also a significant one, since an ordinary perturbation treatment of second order properties from a CI perspective would require a set of $N-1$ excited states, or, equivalently, the inversion of an $N-1 \times N-1$ matrix. Consequently, here also, a systematic reduction of the space is warranted, as this would permit a reasonable approach within the CI framework.

After presenting background material on the partitioning technique, perturbation theory, and Padé approximants in the next two chapters, the general formulation of perturbation theory and the reduced partitioning procedure will be developed in Chapter IV, with the discussion of excited states following in Chapter V. In the final chapter, numerical applications to the H_2 molecule and the HeI^+ molecular ion will be described.

CHAPTER II. THE PARTITIONING TECHNIQUE AND PERTURBATION THEORY

The partitioning technique, as developed by Löwdin [14,15], has been shown to be of primary importance in elucidating the inter-relationship between perturbation theory and the eigenvalue equation as well as leading to a theory of upper and lower bounds to energy eigenvalues. In the following, the partitioning technique will be briefly presented from which the resolvent formulation of perturbation theory will arise as a logical consequence. In addition, the important idea of inner and outer projection will be described in the last section.

II.1. The Partitioning Technique

The paramount objective of molecular quantum mechanics is the solution of the stationary state Schrödinger equation,

$$H\Psi = E\Psi , \quad (2.1)$$

where H is the Hamiltonian for the system and Ψ an eigenfunction. If we assume an expansion for Ψ in terms of an orthonormal n -particle basis $|ff\rangle = |f_1 f_2 \dots f_n \dots\rangle$, then instead of (2.1) we obtain the matrix form of the eigenvalue problem

$$H\mathbf{C} = \mathbf{E}\mathbf{C} \quad (2.2)$$

where the expansion coefficients for each state are grouped into $\mathbf{C} = (\mathbf{c}_0 \mathbf{c}_1 \dots \mathbf{c}_n \dots)$. If \mathbf{f} is complete, then (2.1) and (2.2) are equivalent. For a finite basis, assumed to be of the order $N-1$, the eigenvalues

$(E_0 E_1 \dots E_n)$ are respectively upper bounds to the exact solutions. By defining $M = H - \mathcal{E} \cdot I$, (2.2) can be written

$$M\psi = 0 \quad (2.3)$$

for $\psi = \langle ff | M | ff \rangle$.

If we now consider $|ff\rangle$ to be partitioned into two subsets $|a\rangle$ and $|b\rangle$ of order n_a and n_b , respectively, then it follows that

$$M = \begin{bmatrix} M_{aa} & M_{ab} \\ M_{ba} & M_{bb} \end{bmatrix} \quad (2.4a)$$

and

$$\psi = \begin{bmatrix} \psi_{a0} & \psi_{a1} \dots \psi_{an} \\ \psi_{b0} & \psi_{b1} \dots \psi_{bn} \end{bmatrix} \quad (2.4b)$$

From (2.3) and (2.4), we have the system of equations

$$M_{aa} \psi_{ai} + M_{ab} \psi_{bi} = 0_a \quad (2.5a)$$

$$M_{ba} \psi_{ai} + M_{bb} \psi_{bi} = 0_b \quad (2.5b)$$

for each i . By solving (2.5b) for ψ_{bi} , one obtains

$$\psi_{bi} = -M_{bb}^{-1} M_{ba} \psi_{ai}, \quad (2.6)$$

as long as M_{bb}^{-1} exists. The substitution of (2.6) into (2.5a) gives the expression

$$(M_{aa} - M_{ab} M_{bb}^{-1} M_{ba}) \psi_{ai} = 0_a, \quad (2.7)$$

or, equivalently, if written in terms of the Hamiltonian matrix,

$$[(H_{aa} - \mathcal{E} I_{aa}) + H_{ab} (\mathcal{E} I_{bb} - H_{bb})^{-1} H_{ba}] \psi_{ai} = 0_a. \quad (2.8)$$

Consequently, by defining the modified Hamiltonian matrix

$$\bar{H}_{aa} = H_{aa} + H_{ab} (\varepsilon I_{bb} - H_{bb})^{-1} H_{ba}, \quad (2.9)$$

we have the altered eigenvalue problem

$$\bar{H}_{aa} \psi_{ai} = \psi_{ai} \varepsilon_i, \quad (2.10)$$

which yields n_a roots for $\varepsilon = \varepsilon_i = E_i$.

The choice of partitioning is completely arbitrary, and, in the course of this work, various choices will be made. For the present, however, let us restrict ourselves to the case when $n_a = 1$, and call this single function $f_o = \varphi_o$, with $C_o \neq 0$. From (2.10), we have

$$\varepsilon_o = \bar{H}_{oo} = H_{oo} + H_{ob} (\varepsilon I_{bb} - H_{bb})^{-1} H_{bo}, \quad (2.11a)$$

or, with $E_o^0 = H_{oo} = \langle \varphi_o | H | \varphi_o \rangle$,

$$\varepsilon_o = E_o^0 + \langle \varphi_o | H | b \rangle (\varepsilon I_{bb} - \langle b | H | b \rangle)^{-1} \langle b | H | \varphi_o \rangle. \quad (2.11b)$$

This implicit function of ε , (2.11), is the "bracketing function" which is crucial to the theory of upper and lower bounds [15]. The terminology comes from the fact that for any $\varepsilon > E_o$, then $\varepsilon_o < E_o$ which allows one to "bracket" the eigenvalue E_o .

The connection between (2.11) and Brillouin-Wigner [16] perturbation theory could be obtained directly by separating the inverse matrix into its diagonal and non-diagonal parts and then expanding with respect to the latter by means of the relation

$$(A - B)^{-1} = A^{-1} - A^{-1} B A^{-1} + A^{-1} B A^{-1} B A^{-1} - \dots \quad (2.12)$$

The relationship with perturbation theory, however, is more conveniently presented in terms of the operator formulation of the partitioning technique.

II.2. The Partitioning Technique in Operator Form

In order to discuss the operator formulation it is necessary to introduce the projector

$$\sigma = |\varphi_0\rangle\langle\varphi_0| \quad (2.13)$$

which projects the reference function φ_0 from the space of interest. σ is idempotent and self-adjoint, i.e.,

$$\sigma^2 = \sigma, \quad \sigma^\dagger = \sigma, \quad (2.14)$$

where σ^\dagger indicates the Hermitian adjoint. In addition, the projector for the "orthogonal complement" to φ_0 , the space previously spanned by $|\psi\rangle$, is

$$P = 1 - \sigma, \quad (2.15)$$

with $P^2 = P$ and $P^\dagger = P$ as before. Also, since

$$\sigma P = P\sigma = 0, \quad (2.16)$$

the projectors σ and P are said to be "mutually exclusive."

In order to treat (2.1) we shall also need the definition

$$\bar{H} = P H P \quad (2.17)$$

which is an "outer projection" of H with respect to the subspace of P (see II.5). The outer projected Hamiltonian (2.17) satisfies the relation

$$\bar{H}\bar{\Psi}_k = \bar{E}_k\bar{\Psi}_k \quad (2.18)$$

where $\bar{\Psi}_k$ and \bar{E}_k are its eigenfunctions and eigenvalues, respectively.

The reference function φ_0 is also a trivial eigenfunction of \bar{H} with eigenvalue zero. All the other eigenfunctions $\{\bar{\Psi}_k\}$ are restricted to be in the subspace of P . From a simple theorem about outer projections, it

is easy to see that $\bar{E}_k \geq E_k$ in order, where $\{E_k\}$ are the eigenvalues of H for the eigenfunctions Ψ_k [15(PTX)].

Let us now introduce the complex variable, ε , and subtract ε from both sides of (2.18) to give

$$(\varepsilon - \bar{H})\bar{\Psi}_k = (\varepsilon - \bar{E}_k)\bar{\Psi}_k . \quad (2.19)$$

Using the fact that $P\bar{\Psi}_k = \bar{\Psi}_k$, and defining the "reduced resolvent,"

$$T = (\varepsilon - \bar{H})^{-1}P , \quad (2.20)$$

it follows that (2.19) becomes

$$T\bar{\Psi}_k = (\varepsilon - \bar{E}_k)^{-1}\bar{\Psi}_k , \quad (2.21)$$

demonstrating that T and \bar{H} have the same eigenfunctions. Therefore, the eigenvalue problem (2.21) is completely equivalent to (2.18), but T also contains the complex variable ε which adds to the flexibility of the problem. It can be easily shown that

$$T = TP = PT = PTP , \quad (2.22)$$

and from (2.20), one obtains the useful relationship

$$P(\varepsilon - H)T = P , \quad (2.23)$$

which constitutes a connecting link with the eigenvalue problem (2.1).

Let us now define a "trial" wavefunction Ψ_ε which is dependent on the variable ε by

$$\Psi_\varepsilon = \varphi_0 + T_\varepsilon H \varphi_0 . \quad (2.24)$$

Since $T_\varepsilon H \varphi_0$ is in the orthogonal complement to φ_0 , Ψ_ε is composed of the reference function and a term from the complementary space. With the

definition (2.24) we also have the "intermediate" normalization condition satisfied, that is

$$\langle \varphi_o | \Psi_{\mathcal{E}} \rangle = \langle \varphi_o | \varphi_o \rangle = 1 . \quad (2.25)$$

The intermediate normalization is a convenient choice to make since it is applicable to both the discrete and continuous parts of the spectrum. As long as the normalization integral $\langle \Psi_{\mathcal{E}} | \Psi_{\mathcal{E}} \rangle$ exists, $\Psi_{\mathcal{E}}$ is said to belong to the discrete part. Using (2.23) it is seen that the trial wavefunction satisfies the algebraic identity

$$P(H - \mathcal{E})\Psi_{\mathcal{E}} = 0 \quad (2.26)$$

for all the values of \mathcal{E} . Also, by using the fact that $G + P = 1$, we have

$$(H - \mathcal{E})\Psi_{\mathcal{E}} = \varphi_o(\mathcal{E}_o - \mathcal{E}) \quad (2.27)$$

where \mathcal{E}_o is defined as

$$\mathcal{E}_o = \langle \varphi_o | H | \Psi_{\mathcal{E}} \rangle = \langle \varphi_o | H + HII | \varphi_o \rangle . \quad (2.28)$$

Expression (2.28) is the operator form of the "bracketing function" that is equivalent to (2.11). Hence $\Psi_{\mathcal{E}}$ satisfies an inhomogeneous equation which becomes identical to the eigenfunction in (2.1) when $\mathcal{E}_o = \mathcal{E} = E_o$. The trial function (2.24) and the expression (2.28) are essential to the following treatment of Brillouin-Wigner and Rayleigh-Schrödinger perturbation theory.

II.3. Brillouin-Wigner Perturbation Theory [16]

Assume that the Hamiltonian for the system can be separated such that

$$H = H_o + V, \quad (2.29)$$

where V is a perturbation which may or may not be small. Then if the

previously defined reference function φ_0 is assumed to be the unperturbed solution with Ψ the eigenfunction to the entire Hamiltonian, one may arbitrarily define

$$\Psi = W\varphi_0 \quad (2.30)$$

where W , often termed the "wave operator," takes φ_0 into Ψ . Again, as throughout this paper, we shall require the intermediate normalization

$$\langle \varphi_0 | \Psi \rangle = 1. \quad (2.31)$$

By multiplying the eigenvalue equation (2.1) by φ_0^* (the complex conjugate of φ_0) and integrating we have

$$E = \langle \varphi_0 | H | \Psi \rangle = \langle \varphi_0 | H_0 + V | \Psi \rangle = E_0^0 + \langle \varphi_0 | V W | \varphi_0 \rangle, \quad (2.32)$$

giving the energy difference between the unperturbed result and the exact solution as an expectation value over φ_0 of the quantity VW . As yet we still have no realization for W ; but, by appealing to the previous definitions (2.24) and (2.28) and using the relationship $H_0\varphi_0 = E_0^0\varphi_0$, we obtain

$$\Psi_E = (1 + T_E V)\varphi_0 \quad (2.33a)$$

and

$$\mathcal{E}_0 = E_0^0 + \langle \varphi_0 | V + V T_E V | \varphi_0 \rangle. \quad (2.33b)$$

For the case of $\mathcal{E} = \mathcal{E}_0 = E_0^0$, we have complete agreement with (2.32) for $W = 1 + T_E V$, which defines the primary formulas of perturbation theory.

In all perturbation treatments, we want to have expressions dependent on H_0 and φ_0 . Thus it is convenient to define the reduced resolvent T_0 associated with H_0 in an analogous manner to (2.20), i.e.,

$$T_0 = (\mathcal{E} - P H_0 P)^{-1} P. \quad (2.34)$$

T_o can be related to T by means of the closed form of the identity (2.12),

$$(A - B)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1}, \quad (2.35)$$

now pertaining to operators rather than matrices. With the separation (2.29) and using (2.35), we find that T and T_o are related by

$$T = T_o + T_o V T, \quad (2.36)$$

allowing (2.32) to be written

$$\mathcal{E}_o = E_o^0 + \langle \varphi_o | V + V T_o V + V T_o V T | \varphi_o \rangle. \quad (2.37)$$

By repeated use of (2.36), an infinite geometric expansion for T_E in the quantity $T_o V$ may be obtained,

$$T_E = \sum_{k=0}^{\infty} (T_o V)^k T_o, \quad (2.38)$$

leading to

$$\Psi_E = \varphi_o + \sum_{k=1}^{\infty} (T_o V)^k \varphi_o, \quad (2.39a)$$

and

$$\mathcal{E}_o = E_o^0 + \langle \varphi_o | V [\sum_{k=0}^{\infty} (T_o V)^k] | \varphi_o \rangle. \quad (2.39b)$$

To further illustrate the connection with the more standard inhomogeneous equation approach to Brillouin-Wigner (BW) perturbation theory, we can define

$$\varphi_k = (T_o V)^k \varphi_o \quad (2.40a)$$

and

$$\mathcal{E}_{k+1} = \langle \varphi_o | V | \varphi_k \rangle; \quad (2.40b)$$

then once $\varepsilon_0 = \varepsilon = E$, (2.39) becomes

$$\Psi_E = \varphi_0 + \sum_{k=1}^{\infty} \varphi_k. \quad (2.41)$$

From the recursion formula,

$$\varphi_{k+1} = (T_0 V) \varphi_k, \quad (2.42)$$

and (2.23), we have the set of inhomogeneous equations

$$(E - H_0) \varphi_{k+1} = V \varphi_k - \varepsilon_{k+1} \varphi_0. \quad (2.43)$$

Consequently, as long as it is possible to solve the equations (2.43), one can construct all the $\{\varphi_k\}$ from the unperturbed solution φ_0 and the successive evaluation of the ε_{k+1} 's.

II.4. Rayleigh-Schrödinger Perturbation Theory [17]

In most applications of perturbation theory, the Rayleigh-Schrödinger (RS) type has been used since it generally has better convergence properties than the BW variety, as well as not yielding equations that are implicit functions of the energy, ε . This latter simplification is a consequence of the unperturbed energy appearing in the RS reduced resolvent rather than the ε as occurs in T_0 . Instead of expanding the fundamental resolvent T_ε in terms of V as before, we can instead consider

$$V' = V - (\varepsilon - E_0^\alpha) \quad (2.44)$$

as the quantity to be used in the expansion. If so, from (2.20) and (2.29) we find that

$$T = R_0 + R_0 V' T \quad (2.45)$$

with the RS reduced resolvent defined as

$$R_o = (E_o^0 - P H_o P)^{-1} P. \quad (2.46)$$

In analogy with the BW case, the resolvent and wave operator become

$$T = \sum_{k=0}^{\infty} (R_o V')^k R_o, \quad (2.47a)$$

and

$$W = 1 + \sum_{k=0}^{\infty} (R_o V')^k R_o V. \quad (2.47b)$$

The connection with the conventional inhomogeneous equations of RS theory is somewhat more complicated than in the BW case due to the remainder terms which arise in each order from the $\mathcal{E} = E$ in (2.44). However, from the definitions (2.30) and (2.32) with (2.47), it follows that

$$\Psi_E = \Phi_o + \sum_{k=1}^{\infty} (R_o V')^k \Phi_o \quad (2.48a)$$

and

$$E = E_o^0 + \sum_{k=0}^{\infty} \langle \Phi_o | V (R_o V')^k | \Phi_o \rangle. \quad (2.48b)$$

If the expressions (2.48) are arranged after powers of V in the normal perturbation series expansions,

$$\Psi_E = \Phi_o + \sum_{k=1}^{\infty} \Phi_k \quad (2.49a)$$

and

$$E = E_o^0 + \sum_{k=1}^{\infty} \epsilon_k, \quad (2.49b)$$

then the lower order terms may be identified as

$$\Phi_1 = R_o V \Phi_o \quad (2.50)$$

$$\Phi_2 = R_o (V - \epsilon_1) R_o V \Phi_o = R_o (V - \epsilon_1) \Phi_1$$

$$\begin{aligned}\Phi_3 &= R_o(V - \epsilon_1)R_o(V - \epsilon_1)R_o V \Phi_o - \epsilon_2 R_o^2 V \Phi_o \\ \vdots &= R_o(V - \epsilon_1)\Phi_2 - \epsilon_2 R_o \Phi_1\end{aligned}$$

and

$$\begin{aligned}\epsilon_1 &= \langle \Phi_o | V | \Phi_o \rangle \\ \epsilon_2 &= \langle \Phi_o | V | \Phi_1 \rangle = \langle \Phi_o | V R_o V | \Phi_o \rangle \\ \epsilon_3 &= \langle \Phi_o | V | \Phi_2 \rangle = \langle \Phi_o | V R_o (V - \epsilon_1) R_o V | \Phi_o \rangle \\ \epsilon_4 &= \langle \Phi_o | V | \Phi_3 \rangle = \langle \Phi_o | V R_o (V - \epsilon_1) R_o (V - \epsilon_1) R_o V | \Phi_o \rangle \\ \vdots &\quad - \epsilon_2 \langle \Phi_o | V R_o^2 V | \Phi_o \rangle\end{aligned}\tag{2.51}$$

Eventually the pattern emerges that the RS perturbation wavefunctions Φ_k can be defined recursively as

$$\Phi_{k+1} = R_o(V - \epsilon_1)\Phi_k - \sum_{\ell=1}^k R_o \epsilon_{k-\ell} \Phi_\ell\tag{2.52}$$

with the RS energies simply

$$\epsilon_{k+1} = \langle \Phi_o | V | \Phi_k \rangle.\tag{2.53}$$

These are the usual working definitions of the RS quantities. Using the intermediate normalization condition, (2.52) is easily framed into the inhomogeneous equation form by means of the simple relation

$$P(E_o^0 - H_o)R_o = R_o(E_o^0 - H_o)P = P\tag{2.54}$$

that is obtained analogously to (2.23).

III.5. Inner and Outer Projections

Another important technique which is often used in connection with perturbation theory and variational studies is the method of "inner" and "outer" projections. Since we shall have several occasions for their use, it will be worthwhile to briefly mention some of their properties.

An outer projection of an operator A is defined as

$$\tilde{A} = \sigma_f^A \sigma_f \quad (2.55)$$

for σ_f a projector characteristic of the space spanned by the basis $|\mathbb{F}\rangle$.

For a basis of either finite or infinite order, σ_f may be written explicitly

$$\sigma_f = |\mathbb{F}\rangle\langle\mathbb{F}|^{-1}\langle\mathbb{F}|, \quad (2.56)$$

which is easily seen to satisfy the relations $\sigma_f^2 = \sigma_f$ and $\sigma_f^\dagger = \sigma_f$. In the event that the infinite basis is complete, then σ_f is simply the identity operator. For a moment let us define A as the Hamiltonian operator for the system and assume that

$$\tilde{\Psi}_k = |\mathbb{F}\rangle \mathbb{C}_k \quad (2.57)$$

is an eigenfunction to \tilde{H} . Then from the eigenvalue relation

$$\tilde{H} \tilde{\Psi}_k = \tilde{E}_k \tilde{\Psi}_k, \quad (2.58)$$

we have

$$|\mathbb{F}\rangle (\langle\mathbb{F}| \mathbb{F})^{-1} H_f \mathbb{C}_k - \tilde{E}_k \mathbb{C}_k) = 0. \quad (2.59)$$

Since it is assumed that the basis $|\mathbb{F}\rangle$ is linearly independent,

$$H_f \mathbb{C}_k = \Delta_f \mathbb{C}_k \tilde{E}_k \quad (2.60)$$

where the metric matrix $\Delta_f = \langle\mathbb{F}| \mathbb{F}\rangle$ was introduced. Therefore, the outer projected Hamiltonian \tilde{H} has an exactly soluble eigenvalue equation, which is identical to (2.1) if $|\mathbb{F}\rangle$ is complete and corresponds to the variational solution if the basis is finite. In the latter case, the set of eigenvalues $\{\tilde{E}_k\}$ are guaranteed to be upper bounds to the exact results

in order, since diagonalization of \mathbb{H}_f insures that each $\tilde{\Psi}_k$ is orthogonal and noninteracting with respect to those of lower energy [18,19]. A Hamiltonian of the form \bar{H} is convenient since it allows one to formulate quantum chemical problems in a consistent fashion even though an approximate solution may be involved.

If F and G are two Hermitian operators which fulfill the condition that

$$\langle X|F|X\rangle > \langle X|G|X\rangle \quad (2.61)$$

for all possible X in the common domain of F and G , then it can be said that

$$F > G \quad . \quad (2.62)$$

Consequently, if $F > 0$, F is positive definite.

Now let us require that the operator A be positive definite. If so, we can define the inner projection of A as

$$A' = A^{\frac{1}{2}} \mathcal{O}_f A^{\frac{1}{2}}. \quad (2.63)$$

Since $0 \leq \mathcal{O}_f \leq 1$, A' satisfies the operator inequality

$$0 \leq A' \leq A. \quad (2.64)$$

By making the transformation

$$|\text{In}\rangle = A^{\frac{1}{2}} |\text{ff}\rangle \quad (2.65a)$$

and

$$|\text{g}\rangle = A^{-\frac{1}{2}} |\text{ff}\rangle, \quad (2.65b)$$

we obtain two alternative forms for A' , namely the Bazley projection [20],

$$A' = \left| \langle h \rangle \langle h | A^{-1} | h \rangle^{-1} \langle h | \right|, \quad (2.66a)$$

and the Aronszajn form [21],

$$A' = A \left| \langle g \rangle \langle g | A | g \rangle^{-1} \langle g | A \right|. \quad (2.66b)$$

The inequality (2.64) can only be concluded for $A > 0$, but expressions (2.66) will still converge to A in the limit of a complete basis, constituting a reasonable approximation even though the operator of interest has no unique sign.

One important example of the use of an inner projection is found in the even order terms of a perturbation series expansion. For instance, in the RS case for $R_o < 0$, from (2.46) and (2.51),

$$\epsilon_2 = \langle \Phi_o | V R_o V | \Phi_o \rangle. \quad (2.67)$$

With $P | h \rangle = | h \rangle$, $A = (-R_o)$, and (2.64), we have $R_o' \geq R_o$ giving

$$\epsilon_2' = \langle \Phi_o | V R_o' V | \Phi_o \rangle \geq \epsilon_2, \quad (2.68)$$

which gives an alternate derivation of the Hylleraas variational principle [22,23]. It can be similarly shown that as long as $\Phi_o, \Phi_1, \dots, \Phi_{n-1}$ are exactly known [24],

$$\epsilon_{2n}' \geq \epsilon_{2n} \quad (2.69)$$

for any n , a result obtained by Scherr and Knight in a different way [25]. (See also [26].)

CHAPTER III. PADÉ APPROXIMANTS

In all applications of perturbation theory one is concerned with various series expansions for functions or energies which quite often converge slowly or not at all. In particular, in modern perturbation theory the restriction that the perturbation be small has been essentially eliminated, often leading to very serious convergence difficulties that require more powerful summation techniques in order to extract useful information. One possible approach to this type of problem that has been shown to be fruitful is the Padé approximant summation procedure [27].

Although all the convergence properties of Padé approximants are not yet known [28,29], in the cases where mathematical justification is available, the range of convergence for a Padé approximant has been found to be vastly superior to that of an ordinary power series. In numerous other practical applications where a rigorous mathematical proof is lacking, the results seem to indicate that the range of convergence is actually greatly in excess of that for the situations in which it has been proved. In the following, the basic concept and definitions will be briefly presented including a convenient relationship with the inner projection technique which will subsequently be of use.

III.1. Derivation and Definitions

Consider a function $f(z)$ which has a formal series expansion

$$f(z) = \sum_{\ell=0}^{\infty} a_{\ell} z^{\ell}. \quad (3.1)$$

The $[N,M]$ Padé approximant to this series is defined as the ratio of two

polynomials $P(z)$ and $Q(z)$

$$[N, M] = P(z)/Q(z), \quad (3.2)$$

which are respectively of degree M and N and coincide with the series expansion (3.1) through the $(M + N)^{th}$ power. By requiring the latter property, the coefficients of $P(z)$ and $Q(z)$ are determined by the condition that

$$f(z)Q(z) - P(z) = O(z^{m+n+1}) \quad (3.3)$$

with $O(z^{m+n+1})$ indicating zero up to terms of order greater than $M + N$. To prove that the ratio (3.2) gives a unique $[N, M]$ approximant assume that another ratio (call it P'/Q') is possible, then

$$f(z)Q(z) - P(z) = O(z^{m+n+1}), \quad (3.4a)$$

$$f(z)Q'(z) - P'(z) = O(z^{n+m+1}). \quad (3.4b)$$

Left multiplying (3.4a) by Q' and (3.4b) by Q , the first terms may be eliminated to give

$$P(z)Q'(z) - Q(z)P'(z) = O(z^{m+n+1}). \quad (3.5)$$

Since PQ' and QP' cannot be of higher order than $M + N$, to that order we have

$$P/Q = P'/Q' \quad (3.6)$$

which establishes the uniqueness of the Padé approximant [30].

If we write

$$P(z) = \sum_{k=0}^M b_k z^k \quad (3.7a)$$

and

$$Q(z) = \sum_{k=0}^N c_k z^k , \quad (3.7b)$$

then by performing the multiplication in (3.3) and equating like powers of z , one gets the set of equations

$$\begin{aligned} a_0 c_0 &= b_0 \\ a_0 c_0 + a_1 c_1 &= b_1 \\ \vdots & \\ a_n c_0 + a_{n-1} c_1 + \dots + a_0 c_n &= b_n \\ \vdots & \\ a_m c_0 + a_{m-1} c_1 + \dots + a_{m-n} c_n &= b_m \\ a_{m+1} c_0 + a_m c_1 + \dots + a_{m-n+1} c_n &= 0 \\ \vdots & \\ a_{m+n} c_0 + a_{m+n-1} c_1 + \dots + a_m c_n &= 0. \end{aligned} \quad (3.8)$$

Since the last N homogeneous equations involve the $N + 1$ unknowns, $\{c_i\}$, there is an infinity of solutions possible; but since the Padé approximant is only determined to within a constant factor, by arbitrarily choosing $c_0 = 1$, the remaining $\{c_i\}$ can be obtained. With these c_i 's the $\{b_i\}$ are determined from the first $N + 1$ equations of (3.5) and thus $Q(z)$ and $P(z)$ are specified. The explicit Padé solution to (3.3) is

$$[N, M] = \frac{\begin{vmatrix} a_{m-n+1} & a_{m-n+2} & \cdots & a_{m-1} \\ a_{m-n+2} & a_{m-n+3} & \cdots & a_{m-2} \\ \vdots & \vdots & & \\ a_m & & & \end{vmatrix}}{\begin{vmatrix} \sum_{j=n}^m a_{j-n} z^j & \sum_{j=n-1}^m a_{j-n+1} z^j & \cdots & \sum_{j=0}^m a_j z^j \\ a_{m-n+1} & a_{m-n+2} & \cdots & a_{m-1} \\ a_{m-n+2} & a_{m-n+3} & \cdots & a_{m-2} \\ \vdots & \vdots & & \\ a_m & a_{m-1} & \cdots & a_{m-n} \\ z^n & z^{n-1} & \cdots & 1 \end{vmatrix}} \quad (3.9)$$

subject to the convention that $a_j = 0$ when $j < 0$, and any sums which have the initial element of higher index than the final element are excluded.

Another form of (3.6) that is particularly convenient for computation is the inner projection form [15(PTX), 31, 32]. (See also III.2.) If we define

$$a_i^\dagger = (a_i a_{i+1} \cdots a_{i+N-1}) \quad (3.10)$$

and

$$\Lambda_i = [a_i a_{i+1} \cdots a_{i+N-1}], \quad (3.11)$$

then for $M = N + j$, (3.8) can be written [33]

$$[N, N+j] = \sum_{\ell=0}^j a_\ell z^\ell + z^{j+1} a_{j+1}^\dagger [\Lambda_{j+1} - z \Lambda_{j+2}]^{-1} a_{j+1}, \quad (3.12)$$

where the poles are given by the solutions of $|\Lambda_{j+1} - z \Lambda_{j+2}| = 0$.

The expression (3.12) could be shown to be equivalent to (3.9) by

standard determinant manipulations, but the equality can be established more efficiently by expanding the inverse in (3.12) by (2.12); then

$$[N, N+j] = \sum_{\ell=0}^j a_\ell z^\ell + z^{j+1} \left\{ \sum_{\ell=0}^{2N-1} a_{j+1}^{-1} A_{j+1}^{-1} (z A_{j+2} A_{j+1}^{-1})^\ell a_{j+1} \right\} \\ + \text{"higher order terms"}. \quad (3.13)$$

If we write

$$A_{j+1}^{-1} = \begin{bmatrix} b_{j+1} \\ b_{j+2} \\ \vdots \\ b_{j+N} \end{bmatrix} \quad (3.14)$$

where each b_j is a row matrix of N elements, then from the properties of the inverse

$$b_k a_\ell = \delta_{k\ell}. \quad (3.15)$$

With the notation $\mathbb{1}_k$ meaning a column matrix with the k^{th} element unity and the others zero, we can establish the relations

$$A_{j+1}^{-1} a_{j+k} = \mathbb{1}_k, \quad 1 \leq k \leq N \quad (3.16a)$$

and

$$A_{j+2} \mathbb{1}_k = a_{j+k+1}, \quad 1 \leq k \leq N, \quad (3.16b)$$

from which

$$(A_{j+2} A_{j+1}^{-1})^\ell a_{j+1} = a_{j+\ell}, \quad (3.17)$$

and

$$A_{j+1}^{-1} (A_{j+2} A_{j+1}^{-1})_{j+1}^\ell \mathbb{1}_\ell. \quad (3.18)$$

From (3.18), (3.13) finally becomes

$$[N, N+j] = \sum_{\ell=0}^{2N+j} a_\ell z^\ell + \text{"higher order terms"} . \quad (3.19)$$

By the uniqueness theorem, it necessarily follows that (3.12) is equivalent to (3.9) [33].

III.2. Connection with Moment Expansions and Inner Projections [33]

Let us consider two operators R and A related by

$$R(z) = (1 - zA(z))^{-1}, \quad (3.20)$$

and the expectation value of R with respect to a normalized function ψ ,

$$f(z) = \langle \psi | R(z) | \psi \rangle. \quad (3.21)$$

If R(z) is expanded by (2.35), it follows that

$$f(z) = \sum_{\ell=0}^{\infty} z^\ell \langle \psi | A^\ell | \psi \rangle. \quad (3.22)$$

Defining $a_\ell = \langle \psi | A^\ell | \psi \rangle$, a_ℓ is the ℓ^{th} moment of A and (3.22) is the moment expansion of $f(z)$. For any basis $|h\rangle$, an inner projected approximation to $f(z)$ can be given in the Bazley [20] form by

$$f(z) \sim \langle \psi | h \rangle \langle h | 1 - zA | h \rangle^{-1} \langle h | \psi \rangle. \quad (3.23)$$

Since the basis $|h\rangle$ is arbitrary, we are at liberty to choose the functions

$$|h\rangle = |\psi A \psi^2 \psi \dots \rangle, \quad (3.24)$$

from which it is found that

$$f(z) \sim a_0^\dagger [A_0 - zA_1]^{-1} a_0. \quad (3.25)$$

The result (3.25) can be identified from (3.12) as the $[N, N-1]$ Padé

approximant. Furthermore, since the expansion of (3.20) can also be written as

$$R(z) = 1 + zA(1 - zA)^{-1}, \quad (3.26)$$

it is possible to give another approximation

$$f(z) \sim a_0 + za_1^\dagger [A_1 - zA_2]^{-1} a_1, \quad (3.27)$$

which is seen to be the $[N, N]$ approximant. Similarly, the higher Pade' approximants,

$$f(z) \sim [N, N+j] = \sum_{\ell=0}^j a_\ell z^\ell + z^{j+1} a_{j+1}^\dagger (A_{j+1} - zA_{j+2})^{-1} a_{j+1}, \quad (3.28)$$

are derived. In the event that R is positive definite, from the inner projection property (2.64) it can be further concluded that

$$[N, N] \geq f(z) \geq [N, N-1]. \quad [33] \quad (3.29)$$

In this case the moment expansion is a series of Stieltjes and many additional conclusions may be reached. For other operators that are not of a definite sign, one can usually only treat (3.28) as an approximation to $f(z)$. However, since in the subsequent discussion, the $[N, N-1]$ Pade' approximant will be identified as a partitioned eigenvalue problem, in this instance, it necessarily gives an upper bound for any finite basis.

CHAPTER IV. THE REDUCED PARTITIONING PROCEDURE

By a consideration of perturbation theory within the configuration interaction (CI) domain, we shall develop a procedure that allows an adequate representation for a CI wavefunction with only a small number of perturbation corrections to an unperturbed reference function. This process, termed the reduced partitioning procedure, leads to several advantages over the more standard approaches, one of which is found in the treatment of the full CI problem. The theory as formulated incorporates variational principles with perturbation methods to yield a scheme that has known convergence properties and one which is also very convenient from the computational viewpoint. In connection with the non-linear summation techniques invoked, a modification of ordinary perturbation theory is obtained from which an alternate set of perturbation corrections to the energy and the wavefunction are defined.

IV.1. Non-Linear Summation of Perturbation Expansions

In order to approximately solve the Schrödinger equation,

$$H\Psi = E\Psi, \quad (4.1)$$

via perturbation theory, one usually assumes a separation of the Hamiltonian operator

$$H = H_0 + \lambda V, \quad (4.2)$$

with H_0 chosen to have an eigenfunction Φ_0 and its associated eigenvalue

\mathcal{E}_o , i.e.,

$$H_o \Phi_o = \mathcal{E}_o \Phi_o. \quad (4.3)$$

Within the Rayleigh-Schrödinger framework, it is then assumed that both Ψ and E can be expanded in terms of an "order" parameter λ ,

$$\Psi = \Phi_o + \lambda \Phi_1 + \lambda^2 \Phi_2 + \dots \quad (4.4a)$$

and

$$E = \mathcal{E}_o + \lambda \epsilon_1 + \lambda^2 \epsilon_2 + \dots \quad (4.4b)$$

By substituting (4.4) into (4.1) and equating the various "orders" of λ , one obtains the $\{\Phi_i\}$ as solutions of the inhomogeneous equations

$$(\mathcal{E}_o - H_o) \Phi_n = (V - \epsilon_1) \Phi_{n-1} - \sum_{\ell=0}^{n-2} \epsilon_{n-\ell} \Phi_\ell, \quad (4.5)$$

with the perturbation energies $\{\epsilon_k\}$ defined as

$$\epsilon_{k+1} = \langle \Phi_o | V | \Phi_k \rangle. \quad (4.6)$$

Recalling the definitions

$$\sigma = |\Phi_o\rangle \langle \Phi_o|^{-1} \langle \Phi_o|, \quad P = 1 - \sigma, \quad (4.7)$$

and

$$R_o = (\mathcal{E}_o - PH_o P)^{-1} P, \quad (4.8)$$

then using $R_o \Phi_o = 0$, (2.54), and the intermediate normalization

$$P |\Phi_k\rangle = |\Phi_k\rangle, \quad (4.9)$$

(4.5) may be conveniently written as

$$\Phi_n = R_o(v - \epsilon_1) \Phi_{n-1} - \sum_{\ell=1}^{n-2} R_o \epsilon_{n-\ell} \Phi_\ell. \quad (4.10)$$

If the expansions (4.4) are to yield a valid solution, they must converge to the exact eigenfunction and eigenvalue of the Hamiltonian, H ; but this may not be fulfilled for many choices of the splitting (4.2). Moreover, it is seldom possible to obtain solutions to (4.10) without resorting to some approximate method, a complication which tends to further obscure the validity of the conventional expansions. To increase the flexibility of the problem, we prefer to consider the more general finite expansions

$$\psi^{(m)} = \Phi_o + \lambda \Phi'_1 + \lambda^2 \Phi'_2 + \dots + \lambda^{m-1} \Phi'_{m-1} + \lambda^m \Phi'_m \quad (4.11a)$$

and

$$\epsilon^{(m)} = \epsilon_o + \lambda \epsilon'_1 + \lambda^2 \epsilon'_2 + \dots + \lambda^m \epsilon'_m + \lambda^{m+1} \epsilon'_{m+1} \quad (4.11b)$$

for

$$\Phi'_k = c_k^{(m)} \Phi_k. \quad (4.12)$$

The factors $\{c_i^{(m)}\}$, which may be dependent on the truncation order m , are at our disposal. From the definitions (4.6) and (4.12), the modified perturbation energies are seen to be

$$\epsilon'_{k+1} = \langle \Phi_o | v | \Phi'_k \rangle = c_k \epsilon'_{k+1}. \quad (4.13)$$

In the event that the $\{c_i^{(m)}\}$ for all i and m are chosen to be unity, standard perturbation theory is regained leading to $\psi^{(m)} = \Psi$ and $\epsilon^{(m)} = E$ at infinite order. In other cases, though, we have an additional facility that permits us to insure that (4.11) constitutes a valid expansion that

will allow $\psi^{(m)}$ and $\mathcal{E}^{(m)}$ to approach ψ and E as m increases even though the ordinary linear perturbation series shows no convergence.

The Brillouin-Wigner quantities with

$$T_o = (\mathcal{E} - PH_o P)^{-1} P, \quad (4.14)$$

are defined as

$$\varphi_{k+1} = T_o(\mathcal{E})V\varphi_k \quad (4.15)$$

and

$$\mathcal{E}_{k+1} = \langle \varphi_o | V | \varphi_k \rangle, \quad (4.16)$$

where

$$\mathcal{E} = \langle \varphi_o | V + VT(\mathcal{E})V | \varphi_o \rangle. \quad (4.17)$$

With the previous modification, we similarly obtain

$$\varphi'_k = a_k^{(m)} \varphi_k \quad (4.18)$$

and

$$\mathcal{E}'_{k+1} = \langle \varphi_o | V | \varphi'_k \rangle = a_k^{(m)} \mathcal{E}_{k+1}. \quad (4.19)$$

Besides the conventional choice of c_i or a_i as unity for all i , we have the option of determining the coefficients from the variational principle, or more generally the Lippmann-Schwinger extremum principle [34], where the functions $\{\Phi_i\}$ or $\{\varphi_i\}$ are chosen as the basis vectors. The variational principle has the important advantage of assuring that the modified energy expansion (4.4b) gives a converging upper bound in any order. Another possibility is to determine the coefficients by a Padé summation of the perturbation energies given by (4.6) or (4.16). This

last alternative can yield numerous different solutions of varying degrees of desirability. In particular, the BW energy expansion, (4.17), for a positive definite perturbation and for ϵ less than the first excited unperturbed state, is a series of Stieltjes. From this it may be concluded that determining the $\{a_i^{(m)}\}$ from the $[N,N]$ or $[N,N-1]$ approximants results in the resultant $\epsilon^{(m)}$ for $m \geq 1$ yielding respectively upper and lower bounds to E [33]. (See III.2.) This last feature can also be related to the Lippmann-Schwinger principle [35].

IV.2. A Model Hamiltonian for a Finite Basis

As it is our intention to study energies and other molecular properties as well as is possible within the n -particle basis $|f\rangle^*$, we are necessarily limited to results that could be obtained from the full CI solution. As has been discussed, this in itself is a worthy objective since essentially all of molecular quantum mechanics makes use of some basis set, and, as such, the whole field can rigorously do no better than a full CI within the selected basis.

Moreover, in the vast majority of practical applications of perturbation theory, especially if higher order corrections than φ_1 are desired, the functions defined by (4.10) or (4.15) must be calculated in an approximate manner, usually by introducing a basis. Once the perturbation quantities are so approximated, another question arises, namely will the properties of these approximate solutions be those of the exact ones, and further, will the approximate ones converge to the exact ones as the basis tends to completeness.

*The basis $|f\rangle = |f_0 f_1 \dots f_n\rangle$ in general contains all the configurations of the correct symmetry which can be constructed from a suitably chosen Slater orbital or some other type of basis set.

In order to suitably dispense with these inconveniences from the onset, we shall explicitly introduce the orthonormal basis $|\mathbb{F}\rangle$ by defining a model Hamiltonian as the outer projection of the correct Hamiltonian for the system, then

$$\mathcal{H} = \sigma_f^H \sigma_f = |\mathbb{F}\rangle \langle \mathbb{F}|_H |\mathbb{F}\rangle \langle \mathbb{F}| \quad (4.20)$$

for

$$\sigma_f = |\mathbb{F}\rangle \langle \mathbb{F}|. \quad (4.21)$$

From the splitting (4.2), we also have

$$\mathcal{H}_o = \sigma_f^H \sigma_o \sigma_f \quad (4.22)$$

and

$$\mathcal{V} = \sigma_f^V \sigma_f. \quad (4.23)$$

If we now partition the basis $|\mathbb{F}\rangle$ into a reference function φ_o and n elements required to be in the orthogonal complement to φ_o ,

$$|\mathbb{F}\rangle = |\varphi_o, h\rangle^*. \quad (4.24)$$

With this partitioning, a representation for the projector of the orthogonal complement to φ_o is

$$P = 1 - |\varphi_o\rangle \langle \varphi_o| = |h\rangle \langle h|. \quad (4.25)$$

A useful relationship between the functions $|h\rangle$ and an outer projected operator $\Omega_f = \sigma_f^O \sigma_f$ is that

$$\langle h | \Omega_f | h \rangle = \langle h | \Omega | h \rangle. \quad (4.26)$$

*The partitioning need not be into one function and the remaining n , but may also involve a linear transformation among the elements, $|\mathbb{F}\rangle$. The subsequent discussion and particularly (4.26) will still follow.

By using the result of (II.5), it immediately follows that the model Hamiltonian (4.20) is exactly soluble, i.e.,

$$\mathcal{H}\Psi_i = E_i \Psi_i, \quad (4.27)$$

for

$$\Psi_i = |\psi\rangle \mathbb{C}_i, \quad \mathbb{C} = (\mathbb{C}_0 \mathbb{C}_1 \mathbb{C}_2 \dots \mathbb{C}_n), \quad (4.28)$$

with \mathbb{C} the matrix of eigenvectors to the Hamiltonian matrix defined by $|\psi\rangle$. In addition, (4.22) defines an \mathcal{H}_o which always has a spectrum of exact unperturbed solutions whether or not H_o itself has a set of eigenfunctions. This last feature eliminates the usual perturbation theory tyranny of limiting one to a very few possible forms for H_o , since \mathcal{H}_o can be used for any form of unperturbed Hamiltonian which appears to be desirable. In the event that one actually has an H_o that yields a spectrum of eigenfunctions,

$$H_o \theta_k = \varepsilon_k^o \theta_k, \quad (4.29)$$

then by choosing $|\psi\rangle = |\theta_0 \theta_1 \dots \theta_n\rangle$, it follows from the definitions that

$$\mathcal{H}_o \theta_k = \varepsilon_k^o \theta_k \quad (4.30)$$

also. Thus, the quantity \mathcal{H}_o can be viewed as another generalization of the ordinary theory.

Relative to the finite basis set and the \mathcal{H}_o of (4.22), the previous expressions for the resolvents R_o and T_o become exactly [24]

$$R_o = |\psi\rangle \langle \psi| \mathcal{E}_o - \mathcal{H}_o |\psi\rangle \langle \psi|^{-1} \langle \psi| = |\psi\rangle R_o \langle \psi| \quad (4.31a)$$

$$T_o = |\psi\rangle \langle \psi| \mathcal{E} - \mathcal{H}_o |\psi\rangle \langle \psi|^{-1} \langle \psi| = |\psi\rangle T_o \langle \psi|, \quad (4.31b)$$

where the relationships $P|\psi\rangle = |\psi\rangle$ and (4.26) were used. If the functions

$|h\rangle$ are eigenfunctions of \mathcal{H}_0 , then the inverse matrix is diagonal giving the more customary forms,

$$R_o = \sum_{k=1}^n \frac{|h_k\rangle\langle h_k|}{\varepsilon_o - \varepsilon_k^o} \quad (4.32a)$$

and

$$T_o = \sum_{k=1}^n \frac{|h_k\rangle\langle h_k|}{\varepsilon - \varepsilon_k^o} \quad . \quad (4.32b)$$

Then, subject to the intermediate normalization, the RS expressions are

$$\Phi_n = |h\rangle \alpha_n, \quad (4.33a)$$

$$\alpha_n = R_o [\langle h|V - \varepsilon_1|h\rangle \alpha_{n-1} - \sum_{\ell=1}^{n-2} \varepsilon_{n-\ell} \alpha_\ell] \quad (4.33b)$$

and

$$\varepsilon_{n+1} = \langle \Phi_o | V | h \rangle \alpha_n; \quad (4.33c)$$

and for the BW,

$$\varphi_n = |h\rangle \beta_n, \quad (4.34a)$$

$$\beta_n = T_o \langle h | V | h \rangle \beta_{n-1} \quad (4.34b)$$

and

$$\varepsilon_{n+1} = \langle \varphi_o | V | h \rangle \beta_n. \quad (4.34c)$$

These expressions (4.33) and (4.34) which follow from the model Hamiltonian (4.20) are identical, of course, to those obtained from the Hylleraas, Scherr, and Knight (HSK) method which yields approximate perturbation quantities for the exact Hamiltonian $H = H_0 + V$ [22, 24, 25]. That is, instead of solving the problem (4.1) approximately, we are defining a Hamiltonian $\mathcal{H} = \mathcal{H}_0 + V$ whose eigenfunctions and eigenvalues can be exactly determined.

Our results, however, are still slightly more general, since our \mathcal{K}_o contains the HSK H_o as a special case.

From (4.33) and (4.34) the elements occurring in perturbation theory can be precisely calculated, and the coefficients of the general expansions (4.11) and (4.18) can be chosen to guarantee convergence to the exact result for \mathcal{K} , ensuring that the orthogonality properties of the eigenfunction to an Hermitian operator are obeyed. Furthermore, as $|\psi\rangle$ becomes complete, the solutions of the model Hamiltonian must approach the correct eigenfunctions and eigenvalues of the problem, and by the variational theorem always from above. Similarly, by a selective choice of the coefficients in (4.11) and (4.18), the energy to any order of solution can also be made an upper bound -- and sometimes a lower bound -- which, contrary to much of modern perturbation theory, allows one to always know the relation between the computed result and experiment.

IV.3. The Fundamental Equations for the Reduction Process

In the previous section we defined a perfectly general \mathcal{K}_o as a multi-dimensional outer projection. We shall have occasion to use a form similar to this in the discussion of excited states, but for the present, another form arising from an additional outer projection relative to φ_o ,

$$\mathcal{K}_o = \sigma \mathcal{K}_o \sigma, \quad \sigma = |\varphi_o\rangle\langle\varphi_o|, \quad (4.35)$$

is found to be an especially convenient choice.* Obviously, this \mathcal{K}_o satisfies (4.3) as well as allowing us to introduce any approximate wavefunction as the reference function in a perturbation formulation. From the partitioning (4.24), the definition (4.22), and

*This particular form for \mathcal{K}_o has been previously discussed by Musher [36] and Epstein and Karl [37].

$$P|\psi\rangle = |\psi\rangle, \quad (4.36)$$

it follows that

$$\mathcal{K}_o = |\varphi_o\rangle\langle\varphi_o| \mathbb{E} \langle \mathbb{E} | H_o | \mathbb{E} \rangle \langle \mathbb{E} | \varphi_o \rangle \langle \varphi_o | = |\varphi_o\rangle \mathcal{E}_o \langle \varphi_o | \quad (4.37)$$

for \mathcal{E}_o the expectation value of the arbitrary reference function φ_o . With this \mathcal{K}_o , we have the definition

$$V = \mathcal{H} - \mathcal{K}_o. \quad (4.38)$$

The disadvantage of this particular \mathcal{K}_o is that it leads to divergent expansions [36]. Thus in order to obtain meaningful results, its use necessitates a non-linear summation procedure to obtain the arbitrary coefficients in (4.12) or (4.18).

Another feature of the \mathcal{K}_o of (4.35) is that it only has a single non-vanishing eigenvalue. This results in the R_o of (4.8) and T_o of (4.14) assuming the simple forms

$$R_o = P/\mathcal{E}_o; \quad T_o = P/\mathcal{E}. \quad (4.39)$$

Restricting ourselves for the present to the simpler BW expressions, in the following, we shall deduce a reduction procedure by appealing to the properties of a moment series expansion and the use of Padé approximants. The final closed expression can be identified as an inner projection and, as such, guarantees an upper bound in all orders. From the defining equations of BW perturbation theory, (2.37) and (2.40), we have

$$\mathcal{E} = \langle \varphi_o | V + VT_o V + VT_o VT_o V + \dots | \varphi_o \rangle \quad (4.40)$$

and

$$\varphi_k = (T_o V) | \varphi_{k-1} \rangle = (T_o V)^k | \varphi_0 \rangle. \quad (4.41)$$

By using (4.38) and (4.39), with $P\mathcal{H}_0 = 0$, we obtain

$$\mathcal{E} = \mathcal{E}_0 + 1/\varepsilon \langle \varphi_0 | \mathcal{H} P \mathcal{H} | \varphi_0 \rangle + 1/\varepsilon^2 \langle \varphi_0 | \mathcal{H} (P \mathcal{H})^2 | \varphi_0 \rangle + \dots \quad (4.42)$$

and

$$\varphi_k = 1/\varepsilon^k (P \mathcal{H})^k | \varphi_0 \rangle. \quad (4.43)$$

Then, with the definitions

$$\begin{aligned} \bar{\varphi}_k &= (P \mathcal{H})^k | \varphi_0 \rangle = | \mathbb{H} \rangle \langle \mathbb{H} | H | \mathbb{H} \rangle^{k-1} \langle \mathbb{H} | H | \varphi_0 \rangle \\ &= | \mathbb{H} \rangle H_{\mathbb{H}\mathbb{H}}^{k-1} | \mathbb{H} \rangle, \end{aligned} \quad k \geq 1, \quad (4.44)$$

and

$$\sigma_{k-1} = \langle \varphi_0 | H | \bar{\varphi}_k \rangle = | \mathbb{H} \rangle H_{\mathbb{H}\mathbb{H}}^{k-1} | \mathbb{H} \rangle, \quad (4.45)$$

the $1/\varepsilon$ moment expansion (4.42) can be written as

$$\mathcal{E} = \mathcal{E}_0 + 1/\varepsilon \sigma_0 + 1/\varepsilon^2 \sigma_1 + \dots, \quad (4.46)$$

where the previously defined BW energies are $\mathcal{E}_k = \sigma_{k-2}/\varepsilon^{k+1}$. Consequently, given the Hamiltonian matrix in the complementary space, $H_{\mathbb{H}\mathbb{H}}$, to construct the quantities (4.44) or (4.45), one only needs to perform various matrix multiplications.

In principle it should be possible to solve (4.46) as an implicit function of ε , and if the convergence is sufficiently rapid, obtain a good answer for the eigenvalue. However, this is impossible for two important reasons. First, it is expected that the series (4.46) should be strongly divergent. For example, for the series to converge the eigenvalues $\{\zeta_i\}$ of $H_{\mathbb{H}\mathbb{H}}$ should be such that $-1 < \zeta_i/\varepsilon \leq +1$ for each ζ_i [38]. These conditions are unlikely to be satisfied for molecular problems. A second complication is that for any truncation of (4.46) there is necessarily a

dependence on the energy zero-point. However, since Pade' approximants are known to have better convergence and existence properties than a linear series, by using such a summation technique to determine the coefficients in the elements of the modified energy expansion given by (4.19), it may be possible to improve the situation. In particular, by applying the $[M, M-1]$ Pade' approximant for $M \leq N$, (4.46) can be written as

$$\varepsilon^{(m)} = \varepsilon_o + 1/\varepsilon^{(m)} [M, M-1], \quad (4.47)$$

where

$$[M, M-1] = \sigma_o^\dagger (\Sigma_o - 1/\varepsilon^{(m)} \Sigma_1)^{-1} \sigma_o, \quad (4.48)$$

and

$$\sigma_i^\dagger = [\sigma_i \sigma_{i+1} \dots \sigma_{i+m-1}], \quad i = 0, 1, 2, \dots, m \quad (4.49)$$

$$\Sigma_k = [\sigma_k \sigma_{k+1} \dots \sigma_{k+m-1}], \quad k = 0, 1. \quad (4.50)$$

With some slight rearrangement we have

$$\varepsilon^{(m)} = \varepsilon_o + \sigma_o^\dagger (\varepsilon^{(m)} \Sigma_o - \Sigma_1)^{-1} \sigma_o. \quad (4.51)$$

This is the fundamental formula for the "reduction" process, but as yet it is not obvious if the two previous objections are eliminated. Consequently, let us now consider the functions $|\bar{\psi}\rangle$ defined in (4.44). If we augment the first m of these functions by φ_o , and then assume that $|\varphi_o \bar{\psi}\rangle$ is a reasonable basis for an approximate solution of (4.1), by using the fact that $\langle \varphi_o | \bar{\psi}_k \rangle = \delta_{ok}$, we obtain the partitioned form for the secular equation equivalent to (4.1) as

$$\varepsilon^{(m)} = \varepsilon_o + H_{o\varphi} (\varepsilon_\varphi^{(m)} A_{\varphi\varphi} - H_{\varphi\varphi})^{-1} H_{\varphi o}, \quad (4.52)$$

where the quantities are defined as

$$\Delta_{\varphi\varphi} = \langle \bar{\psi}^{(m)} | \bar{\psi}^{(m)} \rangle, \quad (4.53a)$$

$$H_{\varphi\varphi} = \langle \bar{\psi}^{(m)} | H | \bar{\psi}^{(m)} \rangle, \quad (4.53b)$$

and $\varepsilon_{\varphi}^{(m)}$ is the eigenvalue obtained as a function of the M basis functions $|\bar{\psi}^{(m)}\rangle$. From (4.44) and (4.45) it also follows that

$$H_{\varphi\varphi} = \begin{bmatrix} \sigma_1 & \sigma_2 & \dots & \sigma_m \\ \sigma_2 & \sigma_3 & & \\ \vdots & & & \\ \sigma_m & \sigma_{m+1} & \dots & \sigma_{2m-1} \end{bmatrix} \quad (4.54)$$

and

$$\Delta_{\varphi\varphi} = \begin{bmatrix} \sigma_0 & \sigma_1 & \dots & \sigma_{m-1} \\ \sigma_1 & \sigma_2 & & \\ \vdots & \vdots & & \\ \sigma_{m-1} & \sigma_m & \dots & \sigma_{2m-2} \end{bmatrix}, \quad (4.55)$$

which with $\varepsilon_{\varphi}^{(m)} = \varepsilon^{(m)}$ allows the identification with (4.51). In addition, the intermediately normalized M^{th} order wavefunctions are given by

$$\psi^{(m)} = \varphi_0 + |\bar{\psi}^{(m)}\rangle (\varepsilon^{(m)} E_0 - E_1)^{-1} \varphi_0. \quad (4.56)$$

Therefore, the summation of (4.46) by means of the $[M, M-1]$ Padé approximant is identical to solving the $M+1 \times M+1$ secular equation equivalent to (4.52). As such, besides elucidating the nature of the improved result given by the Padé summation, $\varepsilon^{(m)}$ for any m is guaranteed to be an upper bound to the full solution; and the result obtained from a variational calculation must be independent of the energy zero-point. In the event that $M = N$, we have simply carried out a linear transformation of the original problem which is necessarily equivalent to the full CI solution.

The expressions (4.51) and (4.56) are also seen to be consistent with the definitions (4.18) and (4.19) where the arbitrary coefficients are obtained via the variational principle. The degree of success of the reduced partitioning procedure, however, depends on the [M,M-1] Padé sum approximating the [N,N-1] result even though $M \ll N$, since then $\varepsilon^{(m)}$ should be very close to $\varepsilon^{(n)} = E$, and hopefully, to a desired numerical accuracy. If so, this should effectively negate the necessity of solving enormous secular equations that would arise in the conventional approach, as well as leading to tractable solutions for further use.

The equivalence of the [M,M-1] Padé approximant and the variational solution in the basis $|\phi_0 \bar{\psi}^{(m)}\rangle$ could have been concluded in a slightly different fashion from the deduction of (III.2). This result shows that a moment expansion of which (4.46) is an example can be related to an inner projection which in turn is identified as an [M,M-1] Padé approximant. The fact that it is an upper bound follows from T_0 being a negative definite operator.

Having given the basic equations within the framework of BW theory, it may be asked what happens in the RS case. Recalling the recursion expressions (4.33) and (4.34), and by using (4.26), (4.36), (4.38), and (4.39), we have

$$\bar{\phi}_k = |\hbar\rangle \alpha_k ; \quad \alpha_k = 1/E_0 [H_{hh} \alpha_{k-1} - \sum_{\ell=1}^{k-2} \epsilon_{k-\ell} \alpha_\ell] \quad (4.57)$$

and

$$\bar{\phi}_k = |\hbar\rangle \bar{\beta}_k ; \quad \bar{\beta}_k = H_{hh} \bar{\beta}_{k-1} , \quad (4.58)$$

where (4.58) is slightly modified from (4.34) to be consistent with (4.44). From these two relations, it can be seen that

$$\begin{aligned}\alpha_1 &= 1/\varepsilon_o \bar{\beta}_1 & (4.59) \\ \alpha_2 &= 1/\varepsilon_o^2 H_{hh} \bar{\beta}_1 = 1/\varepsilon_o^2 \bar{\beta}_2 \\ \alpha_3 &= 1/\varepsilon_o^3 H_{hh} \bar{\beta}_2 - \epsilon_2 / \varepsilon_o \bar{\beta}_1 \\ &\vdots\end{aligned}$$

This can be written

$$A = B T \quad (4.60)$$

for $A = (\alpha_1 \alpha_2 \dots \alpha_m)$, $B = (\bar{\beta}_1 \bar{\beta}_2 \dots \bar{\beta}_m)$; and, where it is also observed that T is of upper triangular form. The latter fact ensures that for any order m , the variational solution relative to either basis set is the same; that is,

$$\psi^{(m)} = \varphi_o + |\psi^{(m)}\rangle \alpha^{(m)} = \varphi_o + |\psi^{(m)}\rangle a^{(m)} \quad (4.61)$$

and

$$\epsilon^{(m)} = \langle \psi^{(m)} | H | \psi^{(m)} \rangle. \quad (4.62)$$

Although the variational solutions in the BW and RS bases are equivalent, a slightly different result could have been obtained in the latter case if the coefficients $\{c_i^{(m)}\}$ were determined by applying a $[M, M-1]$ Padé approximant to the series (4.4b). This result is usually only slightly different from the variational solution [24].

IV.4. The Method of Steepest Descent and the Geometric Sumrule

In ordinary perturbation theory, the form of the inhomogeneous equations for the perturbation functions is dictated by the requirement that (4.4) or (4.40) is convergence. In the case of a divergence expansion, it

is not necessarily obvious why the set of perturbation functions should still be one which rapidly exhausts the space of interest, as is necessary for the reduction process to be effective. To attempt some justification of this point, we appeal to the method of "steepest descent" * as employed by McWeeny [40] in connection with SCF studies.

Given an arbitrary trial function, $\varphi = \varphi_0 + \delta\varphi$, where it is assumed that $\langle \varphi_0 | \delta\varphi \rangle = 0$ and $\varepsilon_0 = \langle \varphi_0 | H | \varphi_0 \rangle$, the energy increment due to the first order variation of the wavefunction is

$$\delta\varepsilon = \frac{\langle \delta\varphi | H - \varepsilon_0 | \varphi_0 \rangle + \langle \varphi_0 | H - \varepsilon_0 | \delta\varphi \rangle + \langle \delta\varphi | H - \varepsilon_0 | \delta\varphi \rangle}{\langle \varphi_0 | \varphi_0 \rangle + \langle \delta\varphi | \delta\varphi \rangle}. \quad (4.63)$$

If $| h \rangle$ constitutes a basis in the orthogonal complement to φ_0 , we can assume an expansion for $\delta\varphi$ of the form

$$| \delta\varphi \rangle = | h \rangle \epsilon_1 \quad (4.64)$$

for ϵ_1 the expansion coefficients. If we neglect the second order term, $\langle \delta\varphi | H - \varepsilon_0 | \delta\varphi \rangle$, and ask for the direction of the vector in the linear space corresponding to the maximum possible change in $\delta\varepsilon$, we obtain

$$\delta(\delta\varepsilon) = 0 = \delta\epsilon_1 + [\langle h | H - \varepsilon_0 | h \rangle - \delta\varepsilon \epsilon_1], \quad (4.65)$$

or, that ϵ_1 is proportional to $\langle h | H | \varphi_0 \rangle = b$ as defined previously. Consequently, the basic functional form for the first order correction should be

$$\bar{\varphi}_1 = | h \rangle b. \quad (4.66)$$

The optimum amount of $\bar{\varphi}_1$ to mix with φ_0 , that is the step size η , can be

*Although philosophically similar, the terminology "steepest descent" usually pertains to a different application originally discussed by Riemann and Debye [39].

determined by including the ignored second order term and requiring that

$$\frac{\partial \mathcal{E}(\eta)}{\partial \eta} = 0 , \quad (4.67)$$

for

$$\mathcal{E}(\eta) = \langle \varphi_0 + \eta \bar{\varphi}_1 | H | \varphi_0 + \eta \bar{\varphi}_1 \rangle / \langle \varphi_0 + \eta \bar{\varphi}_1 | \varphi_0 + \eta \bar{\varphi}_1 \rangle . \quad (4.68)$$

Then we have the result that

$$\bar{\eta} = (\mathcal{E}(\bar{\eta})\sigma_0 - \sigma_1)^{-1}\sigma_0 \quad (4.69)$$

where (4.66) and the definitions (4.45) were used. This result gives the first order equations obtained via the steepest descent procedure as

$$\varepsilon^{(1)} = \varepsilon_0 + \sigma_0(\mathcal{E}^{(1)}\sigma_0 - \sigma_1)^{-1}\sigma_0 \quad (4.70)$$

and

$$\psi^{(1)} = \varphi_0 + \bar{\varphi}_1(\mathcal{E}^{(1)}\sigma_0 - \sigma_1)^{-1}\sigma_0 . \quad (4.71)$$

These two equations can be identified as those corresponding to (4.51) and (4.56) for $n = 1$, or the $[1,0]$ Padé approximant solutions.

We can similarly justify the higher order expressions by considering the function (4.71) as a new reference function and repeating the procedure where now $\langle \varphi_0 | \delta\psi \rangle = 0$, but $\langle \psi^{(1)} | \delta\psi \rangle \neq 0$. If so, by again excluding the second order term, we have

$$\begin{aligned} \delta\mathcal{E} &= \frac{1}{\langle \psi^{(1)} | \psi^{(1)} \rangle} \{ \langle \delta\psi | H - \mathcal{E}^{(1)} | \psi^{(1)} \rangle + \langle \psi^{(1)} | H - \mathcal{E}^{(1)} | \delta\psi^{(1)} \rangle \\ &\quad - \delta\mathcal{E}[\langle \delta\psi | \delta\psi \rangle + \langle \delta\psi | \psi \rangle + \langle \psi | \delta\psi \rangle] \}. \end{aligned} \quad (4.72)$$

By assuming $\delta\psi = | h \rangle e_2$, we find that

$$c_2 \sim H_{hh} b + \mu b \quad (4.73)$$

where the constant μ indicates a mixing of some amount of b . Thus, the next function which should be employed to augment $\bar{\varphi}_1$ is exactly the $\bar{\varphi}_2$ of (4.44). By a continuation of this argument, the remaining $\bar{\varphi}$'s can be similarly justified, thereby providing an alternative explanation from a source external to perturbation theory for the special forms of (4.44).

The optimum step size subject to φ_0 , $\bar{\varphi}_1$, and $\bar{\varphi}_2$ is obtained from a form $\psi^{(2)} = \varphi_0 + \eta_1 \bar{\varphi}_1 + \eta_2 \bar{\varphi}_2$, where η_1 and η_2 are variationally determined.

This, of course, is the [2,1] Padé approximant of (4.51) and (4.56).

Another approximation could have been obtained by assuming the $\psi^{(1)}$ of (4.71) to be a new reference function and asking for the steepest change in the first order energy relative to this unperturbed solution. This type of procedure has been employed in a RS framework by Dalgarno and Stewart [41] and Hirschfelder [42,43]. For the BW case, compare also Goldhammer and Feenberg [44] and Young, et al. [45].

Besides being directly derivable from the steepest descent procedure, the first order solutions (4.70) and (4.71) have some other properties which make them particularly significant. For example, their use corresponds to the application of a geometric sum rule to a series, in this case (4.46), and hence is often referred to as the "geometric" approximation. Despite the fact that such an approximation is seldom warranted by the actual terms in the series, it has been used with a great deal of success in a variety of applications [46,47,48,49]. This approximation -- recognized as the [1,0] approximant -- has been previously justified on a variational basis [50], as in the present case, where it rigorously corresponds to a variational solution corrected through the first order in the wavefunction.

As should be expected from the connection with the method of steepest descent, this first order solution gives a substantial improvement over the unperturbed result, yielding, in the cases we have studied, about sixty to seventy per cent of the possible energy improvement that may be gained from the given basis set. Furthermore, the geometric solution also has an important simplifying feature when the reference function is chosen to be the two most common types, namely the Hartree-Fock-Roothaan (HFR) or the CI variety. Let us consider the first case where φ_o is chosen to be an HFR wavefunction and assume the remainder of the configurations are generated by the substitution of a "virtual" orbital. If we partition $|h\rangle$ such that the single, double, etc. excitations are explicitly introduced,

$$|h\rangle = |h^{(1)} h^{(2)} \dots h^{(n)}\rangle, \quad (4.74)$$

then from Brillouin's theorem it follows that

$$\sigma_o = \langle \varphi_o | H | h^{(2)} \rangle \langle h^{(2)} | H | \varphi_o \rangle \quad (4.75)$$

and

$$\sigma_1 = \langle \varphi_o | H | h^{(2)} \rangle \langle h^{(2)} | H | h^{(2)} \rangle \langle h^{(2)} | H | \varphi_o \rangle. \quad (4.76)$$

Consequently, the [1,0] solutions (4.70) and (4.71), relative to an HFR function, depend solely on the double excitations. The CI reference function will be discussed in the next chapter where excited states are considered.

CHAPTER V. EXCITED STATES

The subject of excited states is an important one in quantum chemistry, since it is an area where a great deal of information, which is not readily accessible to experiment, can be obtained from theoretical consideration. Compared to ground states, however, the difficulties associated with an adequate determination of excited states are proportionately compounded. In particular, it is expected that the higher excitations that one usually neglects in a CI wavefunction will be of more significance for excited states, from which it would appear that a reduction process should be of even more importance. In the following, a super-matrix formulation for excited states, which is a generalization of the previous theory, will be presented. From a consideration of these equations, a modification is made which allows a selective treatment of the specific state of interest. This last result will also suggest a conceptually appealing approximate treatment. As in the Hartree-Fock case, the first order solution for excited states is found to have a computationally convenient property.

V.1. The Super-Matrix Formulation

In order to progress to excited states, it is necessary to consider a multi-dimensional reference space^{*} $|x\rangle = |x_0 x_1 \dots x_p\rangle$, which has the characteristic projector

*The multi-dimensional partitioning technique has been employed in lower bound studies by Choi [51] and Wilson [52], where the reference space consisted of a degenerate set of functions.

$$\sigma_X = |\psi\rangle\langle\psi|. \quad (5.1)$$

Within the space spanned by $|\psi\rangle$, for

$$|\psi\rangle = |\psi, h\rangle, \quad (5.2)$$

σ_f is the identity operator, thus

$$P = \sigma_f - \sigma_X = |\psi\rangle\langle\psi|. \quad (5.3)$$

Consider an eigenvalue equation,

$$\mathcal{H}\Psi_i = \Psi_i E_i \quad (5.4)$$

for some state Ψ_i with eigenvalue E_i . Using the property that

$$\mathcal{H} = \sigma_f \mathcal{H} = \mathcal{H} \sigma_f = \sigma_f \mathcal{H} \sigma_f = \sigma_f^H \sigma_f, \quad (5.5)$$

(5.4) can be rewritten as

$$\mathcal{H}(\sigma_X + P)\Psi_i = \Psi_i E_i. \quad (5.6)$$

From (5.6) the two equations,

$$\sigma_X \mathcal{H} \sigma_X \Psi_i = \sigma_X \Psi_i E_i - \sigma_X \mathcal{H} P \Psi_i \quad (5.7a)$$

and

$$P \mathcal{H} \sigma_X \Psi_i = P E_i \Psi_i - P \mathcal{H} P \Psi_i, \quad (5.7b)$$

are obtained by left multiplication of σ_X and P , respectively. Solving (5.7b) for $P\Psi_i$, it follows that

$$P\Psi_i = T_{E_i} \mathcal{H} \sigma_X \Psi_i, \quad (5.8)$$

where T_{E_i} has the usual definition

$$T_{E_i} = (E_i - P \mathcal{K} P)^{-1} P. \quad (5.9)$$

By substituting the expression (5.8) for $P\psi_i$, (5.7a) becomes

$$\Theta_X (\mathcal{K} + \mathcal{K} T_{E_i} \mathcal{K}) \Theta_X \psi_i = \Theta_X \psi_i E_i, \quad (5.10)$$

or that $\Theta_X \psi_i$ is an eigenfunction to the operator $\Theta_X (\mathcal{K} + \mathcal{K} T_{E_i} \mathcal{K}) \Theta_X$ with eigenvalue E_i . (See Löwdin [15(PTIV)].) Thus, (5.10) is exactly equivalent to (5.4), but a consideration of (5.10) leads us to a perturbation treatment.

With the definition (5.1), the fact that $\psi_i = |ff\rangle c_i$, and the partitioning (5.2), (5.10) in matrix form is

$$|\chi\rangle\langle\chi| \mathcal{K} + \mathcal{K} T_{E_i} \mathcal{K} |\chi\rangle\langle\chi| ff\rangle c_i = |\chi\rangle\langle\chi| ff\rangle c_i E_i. \quad (5.11)$$

For $|\chi\rangle$ linearly independent and $c_i = (\frac{c_{X_i}}{c_{h_i}})$, this requires that

$$\bar{\mathcal{H}}_{XX} c_{X_i} = c_{X_i} E_i. \quad (5.12)$$

This gives an alternative operator derivation of (2.10), which was previously obtained from a direct partitioning of the eigenvalue equation. Obviously, the essential difference between the basic equation for the multi-dimensional case, (5.12), and the earlier example for $p = 0$ is that now we cannot remove c_{X_i} to get a direct expression for the energy E_i . Instead, we must solve a secular equation for an effective Hamiltonian. This consequence ensures that the resultant multiple state solutions will satisfy the requisite orthogonality and noninteracting properties. Since the expression (5.12) is an implicit function of E_i , in practice we must resort to the more general form

$$\bar{\mathcal{H}}_{XX}(\varepsilon) c_X(\varepsilon) = c_X(\varepsilon) \varepsilon, \quad (5.13)$$

where for each choice of \mathcal{E} , an $\bar{\mathbb{H}}_{XX}(\mathcal{E})$ may be constructed, from which diagonalization yields $p + 1$ roots, $\{\mathcal{E}_o \mathcal{E}_1 \dots \mathcal{E}_p\}$. The eigenvalues $\{E_i\}$ are given when $\mathcal{E} = \mathcal{E}_i = E_i$ for some i .

With the principal definition,

$$\bar{\mathbb{H}}_{XX}(\mathcal{E}) = \langle \chi | \mathcal{K} + \mathcal{K} T_{\mathcal{E}} \mathcal{K} | \chi \rangle, \quad (5.14)$$

we can proceed as previously by expanding (5.14) via the relation $T_{\mathcal{E}} = T_o + T_o \mathcal{K} T$, to obtain

$$\bar{\mathbb{H}}_{XX}(\mathcal{E}) = \mathbb{H}_{XX} + \langle \chi | \mathcal{K} T_o \mathcal{K} + \mathcal{K} T_o \mathcal{K} T_o \mathcal{K} + \dots | \chi \rangle. \quad (5.15)$$

With the definition

$$\mathfrak{E}_k(\mathcal{E}) = \langle \chi | \mathcal{K} (T_o \mathcal{K})^{k+1} | \chi \rangle, \quad k \geq 0, \quad (5.16)$$

(5.15) can be written as an expansion in terms of matrices

$$\bar{\mathbb{H}}_{XX}(\mathcal{E}) = \mathbb{H}_{XX}(\mathcal{E}) + \mathfrak{E}_o(\mathcal{E}) + \mathfrak{E}_1(\mathcal{E}) + \mathfrak{E}_2(\mathcal{E}) + \dots \quad (5.17)$$

Now let us assume the set $|\chi\rangle$ are the solutions to an unperturbed problem, that is,

$$\mathcal{K}_o |\chi\rangle = |\chi\rangle \mathfrak{E}_o, \quad (5.18)$$

for \mathfrak{E}_o a diagonal matrix consisting of the unperturbed eigenvalues, $\{\mathcal{E}_o \mathcal{E}_1^o \dots \mathcal{E}_p^o\}$. Then the expression (5.17) becomes the multi-dimensional equivalent of the usual BW perturbation expansions which, by taking to a certain order and then solving (5.12), yields the perturbation corrections for each of the $p + 1$ roots. This type of treatment for degenerate RS perturbation theory was originally given by Van Vleck [53]. (See also Hirschfelder [54].) For a general \mathcal{K}_o we could apply various treatments

to (5.17) to obtain series of non-linear expressions for \bar{H}_{XX} , but in the interest of brevity, we shall first specify a particular unperturbed Hamiltonian.

Analogous to the previous single partitioning example, we shall select for the unperturbed Hamiltonian

$$\mathcal{H}_0 = \sum_{\chi} \mathcal{H}_{\chi}^0 \quad (5.19)$$

which is easily seen to satisfy (5.18). Using the spectral expansion and (5.3) we have, just as before,

$$T_0 = P/E. \quad (5.20)$$

Then from (5.16),

$$\epsilon_k(\epsilon) = 1/\epsilon^{k+1} \psi_k = 1/\epsilon^{k+1} \langle \chi | \mathcal{H}(P\mathcal{H})^{k+1} | \chi \rangle, \quad (5.21)$$

where the $\{\psi_k\}$ are independent of ϵ . By applying the [M,M-1] Padé approximant to (5.17) we obtain to order m the super-matrix expression for the reduced, effective Hamiltonian matrix,

$$\bar{H}_{XX}^{(m)} = T_0 + \Gamma_0^\dagger (\epsilon \Omega_0 - \Omega_1)^{-1} \Gamma_0. \quad (5.22)$$

The super-matrices Γ_ℓ and Ω_ℓ are defined by

$$\Gamma_\ell = [\psi_{\ell+1}^\dagger \psi_{\ell+2}^\dagger \dots \psi_{\ell+m-1}^\dagger], \quad \ell = 0, 1, \dots \quad (5.23a)$$

and

$$\Omega_\ell = [\Gamma_\ell \Gamma_{\ell+1} \dots \Gamma_{\ell+m-1}] \quad (5.23b)$$

$$= \begin{bmatrix} i_\ell & \psi_{\ell+1} & \dots & \psi_{\ell+m-1} \\ \vdots & \ddots & & \vdots \\ \psi_{\ell+m-1} & \dots & & \psi_{\ell+2m-1} \end{bmatrix}, \quad \ell = 0, 1, \dots$$

Inspection of (5.22) indicates that the super-matrix product results in a $p+1 \times p+1$ matrix. A somewhat similar kind of treatment has been given in [15(PTXIV)].

To further illustrate the connection with the previous theory, we can define the perturbation corrections relative to each unperturbed state,

$$\varphi_k^{(i)} = (P\mathcal{H})^k |\chi_i\rangle = |\psi_{hh}\rangle H_{hh}^{k-1} \langle h| H |\chi_i\rangle = |\psi_{hh}\rangle H_{hh}^{k-1} b_i \quad (5.24)$$

which should be compared to (4.44). Also, we have the analogous sigma quantities

$$\sigma_k^{(i)} = \langle \chi_i | \mathcal{H} | \varphi_k^{(i)} \rangle = b_i^\dagger H_{hh}^k b_i \quad (5.25)$$

and for $i \neq j$,

$$\sigma_k^{(i,j)} = \langle \chi_i | \mathcal{H} | \varphi_k^{(j)} \rangle = \langle \chi_j | \mathcal{H} | \varphi_k^{(i)} \rangle = b_i^\dagger H_{hh}^k b_j, \quad (5.26)$$

which are generalizations of (4.45). The terms $\sigma_k^{(i,j)}$ are required to describe the interactions arising from the different states $|\chi\rangle$. With these definitions, the elements of the \mathcal{V}_k matrices are seen to be

$$\mathcal{V}_k = [\bar{\sigma}_k^{(0)} \bar{\sigma}_k^{(1)} \dots \bar{\sigma}_k^{(p)}] \quad (5.27)$$

where the sigmas are grouped into column matrices,

$$\bar{\sigma}_k^{(i)} = \begin{bmatrix} \sigma_k^{(0,i)} \\ \sigma_k^{(1,i)} \\ \vdots \\ \sigma_k^{(p,i)} \end{bmatrix} \quad (5.28)$$

As before, we can identify the basic reduction equation for multiple states, (5.22), as the partitional form of an eigenvalue equation, but

this time relative to the basis $|\chi \varphi^{(0)} \varphi^{(1)} \dots \varphi^{(p)}\rangle$, for

$$\varphi^{(i)} = |\varphi_1^{(i)} \varphi_2^{(i)} \dots \varphi_m^{(i)}\rangle. \quad (5.29)$$

When $p = 0$ and $i = 0$, the fundamental expression, (5.22), reduces to the single dimension case.

V.2. The Selective Reduction Procedure

The advantage of a reduction process based on (5.22) is that all states are treated equivalently, giving a set of $p + 1$ improved states at any order m . This is also a disadvantage, though, since in order to effect this result one must use $(m + 1)(p + 1)$ functions in the computation. In a typical CI calculation, even though we expect the order m to be low, p may be a rather large number. To avoid this difficulty, a modification of this method will be presented, which, requiring only $m + p + 1$ functions, permits a higher order treatment for one specific state.

Although it is possible to obtain the desired result directly from a consideration of the super-matrix equations, for illustrative purposes it is convenient to approach the problem from a different perspective. Therefore, let us consider the possibility of obtaining a wavefunction for each excited state from a perturbation treatment of its corresponding unperturbed solution. Then, as in the single partitioning example, we assume a trial function of the form

$$\psi_{\varepsilon_i} = (1 + T_{\varepsilon} \mathcal{K}) \chi_i, \quad (5.30)$$

consisting of a single element in the reference space $|\chi\rangle$, and a term from the orthogonal complement, $|\eta\rangle$. Contrary to a rigorous single partitioning, ψ_{ε_i} cannot become the correct solution Ψ_i when $\varepsilon = \varepsilon_i = E_i$, as long as Ψ_i contains any contributions from the other elements of $|\chi\rangle$.

However, in the usual application of the theory, the set $|x\rangle$ will be eigenvectors of a relatively small CI problem and will be necessarily mutually orthogonal and noninteracting. The only mixing allowed is found in the Hamiltonian matrix elements involving the remaining $\{x_j\}$ and the terms from the orthogonal complement. Consequently, it could be expected that the contribution to Ψ_i from the additional $\{x_j\}$ is small.

In any event, we can obtain the approximate expression

$$E_i \cong \langle x_i | \mathcal{H} | \Psi_{\mathcal{E}_i} \rangle = \langle x_i | \mathcal{H} + \mathcal{H}^T \mathcal{E} \mathcal{H} | x_i \rangle, \quad (5.31)$$

from which the perturbation functions (5.24) are obtained, and the m^{th} order reduction equations analogous to (4.51) and (4.56) are derived;

$$\mathcal{E}_i^{(m)} \cong \tilde{\mathcal{E}}_i^{(m)} = \mathcal{E}_i^0 + \varphi_o^{\dagger(i)} (\mathcal{E} \mathcal{E}_o^{(i)} - \mathcal{E}_1^{(i)})^{-1} \varphi_o^{(i)}, \quad (5.32a)$$

and

$$\Psi_i^{(m)} \cong \tilde{\Psi}_i^{(m)} = x_i + |\varphi^{(i)}\rangle (\mathcal{E} \mathcal{E}_o^{(i)} - \mathcal{E}_1^{(i)})^{-1} \varphi_o^{(i)}. \quad (5.32b)$$

One way to retain the rigorous upper bound properties of the $p+1$ approximate solutions,* (5.32), is to construct the Hamiltonian and overlap matrices relative to $\{\tilde{\Psi}_0, \tilde{\Psi}_1, \dots, \tilde{\Psi}_p\}$ and to solve the secular problem. Then one regains the noninteracting property required by the separation theorem [18,19]. This variation is not very flexible, though, since the relationship between the $\{x_j\}$ and $\{\varphi_k^{(i)}\}$ is already fixed. A superior alternative can be deduced by proceeding somewhat differently. Within the space of interest, and subject to (5.19), if the expression (5.14) is written in more detail, we have

*The solution corresponding to x_0 , the ground state, is still ensured to be an upper bound although it cannot quite converge to the best solution in the given space $|\text{ff}\rangle = |\chi, \mathbb{M}\rangle$ since some of the x_j are excluded.

$$\bar{H}_{XX} = \ell_o + \begin{bmatrix} H_{oh}Q(\varepsilon) H_{ho} & H_{oh}Q(\varepsilon) H_{h1} & \dots & H_{oh}Q(\varepsilon) H_{hp} \\ H_{1h}Q(\varepsilon) H_{ho} & & & \\ \vdots & & & \\ H_{ph}Q(\varepsilon) H_{ho} & \dots & & H_{ph}Q(\varepsilon) H_{hp} \end{bmatrix} \quad (5.33)$$

for $Q(\varepsilon) = (\varepsilon \mathbb{1}_{hh} - H_{hh})^{-1}$. From this exact expression, (5.32) is seen to correspond to a reduction process applied exclusively to the diagonal elements. Consequently, assuming the perturbation functions so obtained, $\{\varphi^{(0)}, \varphi^{(1)}, \dots, \varphi^{(p)}\}$, are a good set to use in a selective reduction treatment, it follows that we can construct several reduced Hamiltonian matrices, one for each $\varphi^{(i)}$, similar to (5.33), by simply taking the off-diagonal elements into account. By so doing, we get to order m

$$\bar{H}_{XX}^{(i,m)} = \ell_o + p_o^{\dagger(i)} (\varepsilon E_o^{(i)} - E_1^{(i)}) p_o^{(i)}, \quad (5.34)$$

where

$$p_k^{(i)} = [\varphi_k^{(0,i)} \varphi_k^{(1,i)} \dots \varphi_k^{(p,i)}], \quad (5.35)$$

and

$$\varphi_k^{\dagger(i,j)} = [\sigma_k^{(i,j)} \sigma_{k+1}^{(i,j)} \dots \sigma_{k+m-1}^{(i,j)}]. \quad (5.36)$$

This form corresponds to a partitioned eigenvalue problem in the basis $|\chi \varphi^{(i)}\rangle$, thereby allowing complete flexibility in the determination of the variation solution to yield the best possible result. At the same time, the proper noninteracting properties are also guaranteed, ensuring upper bounds for all roots. Since one may construct $p+1$ effective Hamiltonian matrices of the form (5.34), diagonalization of each one yields $(p+1)$ eigenvalues and eigenvectors. From the previous discussion of steepest descent, however, it follows that by choosing i as the particular

state of interest, the i^{th} root should be selectively reduced for $m < n$. When $m = n$, (5.34) for any i becomes exactly equivalent to (5.33) giving $p + 1$ identical secular equations. Besides the rigorous solutions, results obtained from the approximate expressions (5.32) will also be found to be useful.

V.3. The Geometric Formula for a Multi-Dimensional Reference Space

In a practical application of the multi-dimensional reduction procedure, one is interested in a CI problem in the space $|\mathbb{F}\rangle$ partitioned into two subspaces $|\mathbb{s}\rangle$ and $|\mathbb{t}\rangle$. Typically, $|\mathbb{s}\rangle$ is composed of a selection of n_s configurations, neglecting the remaining n_t configurations which are grouped into $|\mathbb{t}\rangle$. The dimension of $|\mathbb{s}\rangle$ is usually chosen to be much larger than the number of roots desired. In such a case, the first order solution, a type of geometric formula, has a convenient simplifying feature.

Let us consider

$$|\mathbb{F}\rangle = |\mathbb{s} \ \mathbb{t}\rangle \quad (5.37)$$

for $|\mathbb{F}\rangle$ an orthonormal set of configurations. The equation we would like to solve is (5.4), or equivalently,

$$\mathbb{H}_f \mathbb{C}_f = \mathbb{C}_f \mathbb{E}. \quad (5.38)$$

In practice, though, we usually must resort to solving a truncated problem within the chosen space, $|\mathbb{s}\rangle$. By so doing, from the definitions (4.20) and (5.37), we have

$$\langle \mathbb{s} | \mathbb{H} | \mathbb{s} \rangle = \langle \mathbb{s} | \mathbb{H} | \mathbb{s} \rangle = \mathbb{H}_{ss}, \quad (5.39)$$

and

$$\mathbb{H}_{ss} \mathbb{D}_s = \mathbb{D}_s \boldsymbol{\epsilon}_o, \quad (5.40)$$

for the $\boldsymbol{\epsilon}_o$ of (5.18) and the n_s possible eigenvectors,

$$\mathbb{D}_s = [\mathbf{d}_o \mathbf{d}_1 \dots \mathbf{d}_{n_s-1}]. \quad (5.41)$$

Since we are only interested in the first $p + 1$ solutions, let us define them as

$$\chi_i = |\$ \rangle \mathbf{d}_i, \quad i = 0, 1, \dots, p \leq n_s. \quad (5.42)$$

These wavefunctions will constitute a multi-dimensional unperturbed reference space to implement a perturbation solution of (5.38). Now we may partition $|\mathbb{H}\rangle$ such that

$$|\mathbb{H}\rangle = |\chi, \bar{\chi}, \mathbb{W}\rangle = |\chi, \mathbb{W}\rangle, \quad (5.43)$$

where $|\bar{\chi}\rangle$ consists of the remaining $n_s - (p + 1)$ functions, which may be obtained from the basis $|$\rangle$. The $|\bar{\chi}\rangle$ are assumed to be orthogonal to $|\chi\rangle$, and an adequate realization would be the other eigenvectors of (5.41), which are necessarily noninteracting with the unperturbed solutions. The following is a consequence of this latter property.

The first order solution to the super-matrix equation, (5.22), is

$$\bar{\mathbb{H}}_{XX}^{(1)} = \boldsymbol{\epsilon}_o + \mathbb{V}_o^\dagger (\mathcal{E} \mathbb{V}_o - \mathbb{V}_1)^{-1} \mathbb{V}_1. \quad (5.44)$$

From the relation (5.27), the important components are the $\{\sigma_k^{(i,j)}\}$ defined in (5.26) where $k = 0$ and $k = 1$. With the definition of \mathbb{H} and by using the eigenfunction properties of the $|\chi\rangle$, it follows that

$$\sigma_o^{(i,j)} = \langle \chi_i | \mathbb{H} | \mathbb{W} \rangle \langle \mathbb{W} | \mathbb{H} | \chi_j \rangle, \quad (5.45)$$

and

$$\sigma_1^{(i,j)} = \langle \chi_i | H | \psi \rangle \langle \psi | H | \psi \rangle \langle \psi | H | \chi_j \rangle, \quad (5.46)$$

for any $i, j = 0, 1, \dots, p$. This has the convenient consequence that in a truncated CI calculation the first order improvement to the $p + 1$ lowest states can be obtained by ignoring the $n_s - (p + 1)$ combinations, $|\bar{\chi}\rangle$, and solely introducing the excluded configurations, $|\psi\rangle$, into the calculation. Since the selective reduction equations (5.34) and the approximate solutions (5.32) are sub-classes of (5.22), this feature is also present for these methods. Although we employed the $n_s - (p + 1)$ remaining eigenvectors of (5.41) in this derivation, in the interest of going to higher orders, it may be mentioned that this is not required since only the orthogonality of the $|\bar{\chi}\rangle$ to the unperturbed solutions, $|\chi\rangle$, is sufficient to ensure the noninteracting property.

When one selects a set of configurations for a CI calculation, it is assumed that those remaining are relatively insignificant. By determining the first order improvement to the initial choice $|\psi\rangle$ by means of (5.45) and (5.46), this supposition can be placed on a more quantitative basis. Its assistance in choosing important configurations is also apparent. Furthermore, since this solution typically accounts for more than fifty per cent of the obtainable energy improvement, it seems that it would be a logical extension of any standard CI computation.

CHAPTER VI. NUMERICAL APPLICATIONS OF THE REDUCED PARTITIONING PROCEDURE

In order to undertake a numerical assessment of the reduced partitioning theory, we made some numerical applications within the configuration interaction framework. Since one of the objectives of the theory is to minimize the difficulties encountered in configuration selection, we were interested in limiting these initial computations to molecules where the full CI wavefunction for a basis of several functions would be easily tractable, since the full CI, being the best answer in the chosen basis, will provide the ultimate result to use in our comparisons. Practically, this limited us to the two-electron systems H_2 and HeH^+ , as the number of initial basis functions allowed by our computer program (18) would lead to a few thousand configurations of the proper symmetry for even LiH .

The calculations that we have performed fall essentially into two categories: ground state studies relative to a single unperturbed solution, and a simultaneous treatment of several states by employing a multi-dimensional reference manifold. The first of these will be illustrated with equilibrium studies of H_2 relative to a simple Coulson molecular orbital wavefunction as the unperturbed state as well as with a Hartree-Fock-Roothaan reference function. In addition, a complete potential curve for the lowest $^1\Sigma^+$ state of HeH^+ subject to a HFR unperturbed solution will be obtained. As an example of the multi-dimensional case, the selective reduction of the $^1\Sigma^+$ excited states for HeH^+ will also be considered, including some numerical justification of the approximate treatment previously described. Although this discussion is self-contained in the

sense that the objective is a theoretical solution, due to the considerable experimental interest in the reactions of metastable He [55,56], these calculations also may be of some interest in connection with this problem.

VI.1. Summary of the Computational Procedure

The computer program implemented for this problem couples the ingredients of a configuration interaction calculation with the perturbation approach described in the previous chapters. The initial basis functions used are the complex form of the Slater type orbitals (STO's) located on each center in the molecule. The complex form ensures that each orbital is an angular momentum eigenstate and thus assists one in constructing n-particle functions of the proper symmetry. From the initial basis of STO's, a secondary basis consisting of an orthogonal set of orbitals is generated. In the calculations presented here, these are either the Schmidt orthogonalized combinations of the STO's or the Hartree-Fock-Roothaan eigenvectors one obtains from solving a preliminary single determinant problem. By associating α or β spin with each member of this secondary basis, a set of one-particle pure spin orbitals is obtained. Due to the symmetry of the STO's, the resultant spin orbitals are specified by the usual σ , π , δ , etc. classification. From the spin orbitals, n-particle anti-symmetrized products are constructed, where the symmetry of the state Σ , Π , Δ , etc. is easily achieved from the spin orbital specifications. For homonuclear diatomics the gerade and ungerade symmetry for the orbitals is also taken into account. By combining these anti-symmetrized products into the proper combinations, spin symmetry is incorporated into the elements of the basis set, as is the plus and minus symmetry arising from reflection in the plane of the molecule for sigma

states. A configuration for a diatomic molecule is thus defined as a pure symmetry function in all respects, as this gives the minimum possible number that must be considered for the problem of interest.

The molecular integrals are computed from a program written by H. H. Michels and altered by J. Kouba to give integrals accurate to at least eight figures. Additional routines from J. Kouba's CI program [57] have been modified by the author for inclusion into the completed program. When Hartree-Fock wavefunctions are required, they are obtained from a program written by H. H. Michels, but with the more accurate integral routines substituted.

It may be worthwhile to briefly summarize the essential steps in the remainder of the program, which is used to effect the perturbation calculations. This section is arranged to calculate CI wavefunctions relative to a given set of configurations, and then make two sub-partitionings in each single run. With each partitioned result constituting one or several unperturbed states, these solutions and the remaining elements of the orthogonal complement are used to construct the perturbation quantities by the theory previously described. In this respect, routines have been written which allow us to work either within the RS or BW framework. Since a variational solution relative to the perturbation basis is desired, this can be accomplished by generating the respective Hamiltonian and overlap matrices from which the eigenvalue equation may be solved. In the BW case, this is easily accomplished solely from the perturbation energies, since from the basic definitions, (2.40), it follows that

$$\langle \varphi_p | v | \varphi_q \rangle = \langle \varphi_p | v(T_o v)^q | \varphi_o \rangle = \langle \varphi_{p+q} | v | \varphi_o \rangle = \epsilon_{p+q+1} \quad (6.1a)$$

and

$$\langle \varphi_p | \varphi_q \rangle = \epsilon_{p+q}. \quad (6.1b)$$

In the RS case, the overlap matrix elements must be calculated from the perturbation corrections (2.52), but from these overlap elements and the perturbation energies (2.53), the Hamiltonian matrix elements are efficiently constructed via a relationship derived by Löwdin [15(PTJX)] giving

$$\langle \Phi_p | V | \Phi_q \rangle = \epsilon_{p+q+1} + \sum_{k=1}^p \sum_{\ell=1}^q \epsilon_{p+q+1-k-\ell} \langle \Phi_k | \Phi_\ell \rangle. \quad (6.2)$$

Within BW theory, an iterative procedure for obtaining the $[M, M-1]$ and $[M, M]$ Padé approximants by means of the solution of the implicit equation, $\mathcal{E} = f(\mathcal{E})$, has also been implemented. This routine incorporates the technique of Newton-Raphson iterations [15(PTI)], a second order process, to increase the rate of convergence. For example, for the $[M, M-1]$ approximant with a single partitioning

$$\mathcal{E} = f(\mathcal{E}) = E_0^{(o)} + \varphi_o^\dagger (\mathcal{E} E_o - E_1)^{-1} \varphi_o, \quad (6.3)$$

from which one readily obtains the quantity

$$f'(\mathcal{E}) = -\varphi_o^\dagger (\mathcal{E} E_o - E_1)^{-1} E_o (\mathcal{E} E_o - E_1)^{-1} \varphi_o, \quad (6.4)$$

which is always negative. From a first order approximation

$$\mathcal{E}^{(1)} = f(\mathcal{E}^{(o)}), \quad (6.5)$$

the improved value,

$$\mathcal{E}^* = \mathcal{E}^{(o)} + [(\mathcal{E}^{(1)} - \mathcal{E}^{(o)}) / (1 - f'(\mathcal{E}^{(o)}))], \quad (6.6)$$

follows from an application of the Newton-Raphson formula to

$$F(\mathcal{E}) = \mathcal{E} - f(\mathcal{E}) = 0. \quad (6.7)$$

Furthermore, the fact that $f'(\mathcal{E}) < 0$ guarantees that for $\mathcal{E}^{(o)} > \mathcal{E}$,

$\varepsilon^{(1)} < \varepsilon$; but ε^* , which can be related to the variational principle [15(PTI)], must always be greater than ε , ensuring convergence from above. This process is continued until two successive ε 's differ by no more than 10^{-10} . Since this method requires repeated inversions of $(\varepsilon E_0 - E_1)$ for the various values of ε , although it may be more desirable in some respects, it is a more time consuming process than the equivalent secular equation solution.

Despite the particular approach used, one encounters computational problems with near linear dependency in the perturbation basis set. In the implicit function process, this complication is manifested in the inversion of the ill-conditioned matrix $(\varepsilon E_0 - E_1)$, and, in the eigenvalue equation, by the removal of the overlap matrix. In the program, this possibility is investigated at each order of calculation by multiplying the matrix with its inverse, or checking the transformed overlap matrix. As the perturbation functions rapidly exhaust the space of interest, a resulting near linear dependency is indicative of having essentially achieved the desired solution, but in the process it is often of interest to gain another one or two figures of accuracy. Consequently, to obtain slightly higher orders of solution than would otherwise be possible, a multi-precision extension written by C. E. Reid [58] for the IBM 360/65 has been incorporated into the program. With this package, multi-precision routines for the Padé approximant analysis, including a matrix inversion subroutine, have been written. The use of these routines depends on the assumption that the pertinent quantities need to be calculated only in double precision (16 figures). Then with these values, the remaining analysis involving the more sensitive steps is accomplished with four words (37 figures) of precision. This process can be trusted up to the point where the linear dependency once again manifests itself,

or when the ultimate effect depends on differences beyond the range of the low precision of the initially calculated quantities. Either of these features becomes apparent in the course of the calculation.

In order to treat multiple states with the selective reduction procedure, the program uses essentially the same computational framework as in the single partitioning case. To illustrate this, let us consider the secular equation approach which requires constructing the "reduced" Hamiltonian and overlap matrices,

$$H_R^{(i)} = \begin{bmatrix} H_{XX} & H_{X\varphi}^{(i)} \\ \hline H_{\varphi X}^{(i)} & H_{\varphi\varphi}^{(i)} \end{bmatrix}, \quad (6.8)$$

and

$$\Delta_R^{(i)} = \begin{bmatrix} \Delta_{XX} & \Delta_{X\varphi}^{(i)} \\ \hline \Delta_{\varphi X}^{(i)} & \Delta_{\varphi\varphi}^{(i)} \end{bmatrix}. \quad (6.9)$$

The partitionings are defined as

$$H_{XX} = E_o, \quad (6.10a)$$

$$H_{X\varphi}^{(i)} = \langle p^{(i)} | \mathcal{K} | \chi \rangle = [\sigma_1^{(i,0)} \sigma_2^{(i,1)} \dots \sigma_m^{(i,p)}], \quad (6.10b)$$

$$H_{\varphi\varphi}^{(i)} = \langle p^{(i)} | \mathcal{K} | p^{(i)} \rangle = E_1^{(i)};$$

and

$$\Delta_{XX} = \langle \chi | \chi \rangle = 1_\chi, \quad (6.11a)$$

$$\Delta_{X\varphi}^{(i)} = \langle p^{(i)} | \chi \rangle = \emptyset, \quad (6.11b)$$

$$\Delta_{\varphi\varphi}^{(i)} = \langle p^{(i)} | p^{(i)} \rangle = E_o^{(i)} \quad (6.11c)$$

for the reference functions, $|\chi\rangle$, and the sigma terms defined in (4.49), (4.50), and (5.26).

Since we have \mathbf{E}_0 from the unperturbed solution, the quantities needed are $H_{\chi\varphi}^{(i)}$, $H_{\varphi\varphi}^{(i)}$, and $\Lambda_{\varphi\varphi}^{(i)}$. To obtain these matrices, it is convenient to begin by calculating the approximate solution, (5.32), for each of the $p + 1$ states. As this is a single partitioning approach, this is achieved with the basic program structure. In particular, the quantities

$$\mathbf{b}_i = \langle \chi_i | \mathcal{K} | h \rangle \quad (6.12)$$

and the perturbation functions

$$\varphi^{(i)} = | h \rangle \beta^{(i)} \quad (6.13)$$

are generated for each i . From these, one also produces the matrices $H_{\varphi\varphi}^{(i)}$ and $\Lambda_{\varphi\varphi}^{(i)}$, which defines the secular equation for the approximate solution, (5.32). By solving this, the approximate energy, $\tilde{\epsilon}_i$, and wavefunction, $\tilde{\psi}_i$, are given as the lowest energy solution. In the process of obtaining each of these approximate results, by simply storing the $\beta^{(i)}$, \mathbf{b}_i , $H_{\varphi\varphi}^{(i)}$, and $\Lambda_{\varphi\varphi}^{(i)}$ externally, all the basic information necessary for the rigorous multiple states solutions is available.

The remaining elements required are the terms $\{\sigma_k^{(i,j)}\}$ which arise from the interaction of the unperturbed states, and are thus unique to the multi-dimensional case. These present no problem, though, since from the definition

$$\sigma_{k-1}^{(i,j)} = \langle \chi_j | \mathcal{K} | \varphi_k^{(i)} \rangle, \quad k \geq 1, \quad (6.14)$$

one immediately obtains

$$\sigma_{k-1}^{(i,j)} = \mathbf{b}_j \beta^{(i)}. \quad (6.15)$$

Following this strategy, the program efficiently constructs the $p + 1$

possible reduced Hamiltonian and overlap matrices, from which the secular solution gives $p + 1$ roots, of which the i^{th} one is selectively reduced. As a bonus the approximate solutions are given, and as our computations will indicate, these are likely to be sufficiently close to constitute a satisfactory treatment for many applications.

To place the computational benefits of the reduced partitioning theory into perspective, it should be noted that in any truncated CI calculation with a basis set, the calculation of the molecular integrals is the predominant time consuming operation, with a considerable amount of time also required for their transformation to a secondary basis. Since both of these steps are necessary despite the number of configurations used in the calculation, it seems logical to get maximum benefit from the chosen basis by using many more configurations than is customary, or even all that are possible. To do this in the standard approach would involve solving an $N \times N$ secular equation for several states, which is, at present, an impossibility for very large N . On the other hand, the reduction process does not require such a diagonalization, but, as will be shown, gives a result in essential agreement with the full solution by only solving low order secular equations. It is still generally necessary to construct all the Hamiltonian matrix elements, but even though this is admittedly a problem for very large N , since these elements are obtainable solely from the transformed molecular integrals and the configurations, the extension is reasonable.

VI.2. Results and Discussion: The Single Partitioning Case*

As a first example of the reduced partitioning procedure (RPP), we considered the hydrogen molecule at the equilibrium internuclear separation

*All the calculations reported herein were done on the IBM 360/65 computer of the University of Florida Computer Center.

$R = 1.40$ a.u. The n -particle basis, $|\varphi_0^{\text{sh}}\rangle$, was generated from an initial set of nine STO's on each atom, $1s(1.20)$, $1s'(1.00)$, $2s(1.16)$, $2p_0(1.71)$, $2p_{\pm 1}(1.71)$, $3d_0(2.20)$, and $3d_{\pm 2}(2.00)$,^{*} which, using Schmidt orthogonalized combinations, led to 34 possible configurations of $^1\Sigma_g^+$ symmetry. As this is the full CI basis, it gives the best answer obtainable for the chosen space. The full CI result is $E = -1.1712787$ a.u.

The principal contribution to the CI wavefunction is the Coulson function configuration^{**}

$$\varphi_0 = \langle 1s_g^{\alpha(1)} 1s_g^{\beta(2)} \rangle \quad (6.16)$$

which constitutes $\approx 98\%$ of the full solution and gives $E = -1.1280771$ a.u. by itself. Choosing this configuration as the reference function, the $\{\bar{\varphi}_k\}$ and the $\{\sigma_i\}$ were calculated from (4.44) and (4.45), and from these quantities the reduction energies were obtained from (4.51) for each m from 1 to 13. These numbers are listed in Table 1.

For comparison, the difference between the unperturbed energy and the full CI result will be defined as

$$\Delta E(i) = \mathcal{E}_i^0 - E_i \quad (6.17)$$

where i indicates the state of interest. If \mathcal{E}_i^0 is a HF energy, then ΔE may be referred to as E_{corr}^f , the correlation energy obtainable within the space $|ff\rangle$, by analogy with the usual definition of correlation energy.

The most dramatic feature illustrated in Table 1 is the anticipated extreme divergence of the sigma series which, lacking a non-linear summation technique, would be useless. However, by applying the $[M,M-1]$ Padé

*The orbital exponents for $n = 2$ were taken from the optimized basis set for $R = 1.40$ a.u. obtained by McLean, et al. [59].

**In fact, Coulson [60] found the best scale factor equal to 1.197 for $R = 0.732$ Å.

TABLE 1

Reduction Energies for H₂ with the Coulson Reference Function^a
(R = 1.40)

M	$\mathcal{E}^{(m)} = \mathcal{E}_0 + 1/\mathcal{E}^{(m)}[M, M-1]$	% E _{corr} ^f	Brillouin-Wigner Sigma Elements
0	-1.1280771	0.00	$\sigma_0 = 0.124303$
1	-1.1582136	69.76	$\sigma_1 = 0.279955$
2	-1.1662436	88.35	$\sigma_2 = 1.594896$
3	-1.1696479	96.23	$\sigma_3 = 1.25 \times 10^1$
4	-1.1707700	98.82	$\sigma_4 = 1.18 \times 10^2$
5	-1.1711049	99.60	$\sigma_5 = 1.22 \times 10^3$
6	-1.1712244	99.87	:
7	-1.1712552	99.95	$\sigma_{25} = 2.23 \times 10^{23}$
8	-1.1712685	99.98	
9	-1.1712741	99.99	
10	-1.1712764	99.99	
11	-1.1712778	100.00	
12	-1.1712782	100.00	
13	-1.1712784	100.00	
33(Full CI)	-1.1712787		

^aEnergies and bond distances are always expressed in atomic units.

approximants to the series, it is seen that the reduction energies converge very rapidly toward the full CI solution [61]. From the connection between the $[M, M-1]$ approximant and the variational treatment, it is more easily seen why this result is found [61]. This is the same situation previously observed in an application of BW perturbation theory to the hydride ion [62], where Padé approximants were also used to sum the series to give meaningful energies.

As has been mentioned, the fact that the perturbation functions $|\varphi_0, \varphi_1, \dots, \varphi_m\rangle$ rapidly exhaust the space of interest causes problems with near linear dependency. In the examples studied, this usually becomes apparent in the ninth or tenth order solution. Consequently, multiple precision was used to obtain the higher order solutions in Table 1, permitting another one or two digits of agreement.

Due to the simplification inherent in the first order solution, it is especially encouraging that this result accounts for almost seventy per cent of the possible energy improvement. For many problems it may be difficult to calculate the higher order terms, thus a dramatic energy decrease from the "geometric" result is a definite asset.

In order to consider a HFR reference function, the previous STO basis set was used in a HFR calculation to give a set of one-electron orbitals as eigenfunctions to the Fock Hamiltonian. From these, another 3/4-configuration CI wavefunction was constructed but with the principal configuration now actually being the HF result. The energy of the HF solution is -1.1335224 a.u., and, of course, the full CI energy is the same as before. For a molecule such as H_2 , where the $1s_g$ STO combination is a good approximation to the $1\sigma_g$ molecular orbital, there cannot really be too much difference between the Coulson function and the HF solution. This is reflected in a comparison of Table 1 and Table 2. The elements of the sigma

TABLE 2

Reduction Energies for H₂ with a Hartree-Fock Reference Function
(R = 1.40)

M	$\epsilon^{(m)} = \epsilon_0 + 1/\epsilon^{(m)}[M, M-1]$	% E _{corr} ^f	Brillouin-Wigner Sigma Elements
0	-1.1335224	0.00	$\sigma_0 = 0.106462$
1	-1.1590861	67.71	$\sigma_1 = 0.243929$
2	-1.1668000	88.14	$\sigma_2 = 1.509162$
3	-1.1696330	95.64	$\sigma_3 = 1.26 \times 10^1$
4	-1.1707910	98.71	$\sigma_4 = 1.22 \times 10^2$
5	-1.1711357	99.62	$\sigma_5 = 1.28 \times 10^3$
6	-1.1712490	99.92	:
7	-1.1712697	99.98	$\sigma_{23} = 1.27 \times 10^{22}$
8	-1.1712763	99.99	
9	-1.1712779	100.00	
10	-1.1712785	100.00	
11	-1.1712786	100.00	
12	-1.1712786	100.00	
33(Full CI)	-1.1712787		

expansion are very similar as are the reduction energies for each order, although the results with the HF reference function do seem to converge a little better in higher order. The energy decrease given by the geometric solution is somewhat better for the Coulson function, which is probably a result of more improvement being possible. It should be observed that better than ninety-five per cent of the possible correlation energy is given by only a third order or four-function expansion.

For the HeH^+ molecular ion a very good HF calculation, including orbital exponent optimization, has been done by Peyerimhoff [63]. For $R = 1.455$ a.u., the optimized basis she obtained is $1s_{\text{He}}(1.37643)$, $1s'_{\text{He}}(3.87107)$, $2s_{\text{He}}(1.54335)$, $2p_o \text{He}(2.64576)$, $2p'_o \text{He}(3.24082)$, $3d_o \text{He}(2.54147)$, $4f_o \text{He}(3.73526)$, $1s_{\text{H}}(1.00949)$, $2s_{\text{H}}(1.18036)$, $2s'_{\text{H}}(2.56229)$, $2p_o \text{H}(1.79089)$, $3d_o \text{H}(2.41228)$, from which, she reports an energy of $E_{\text{HF}} = -2.933126$ a.u. By repeating this HF calculation, we obtained the slightly higher result, $E_{\text{HF}} = -2.933072$ a.u., which we expect is due to our more accurate integral computation. By augmenting this sigma basis with $2p_{\pm 1} \text{He}(2.868)$, $2p_{\pm 1} \text{H}(0.827)$, and $3d_{\pm 2} \text{He}(2.686)$,^{*} a full CI wavefunction consisting of 82 configurations of ${}^1\Sigma^+$ symmetry was obtained. With the HF solutions as the reference function, the series elements and the reduction energies are as given in Table 3. It appears the sigma series is even somewhat more divergent than in the H_2 case which might be expected from the convergence criterion given in (IV.3), but may also be dependent on the adaptability of the basis set, or the number of configurations. In keeping with the greater extent of divergence, the convergence of the reduction energies is slightly worse than before, even though the third order solution still accounts for more than ninety per cent of the available correlation energy. The geometric result gives about fifty-nine per cent.

*This STO basis will be referred to as basis I.

TABLE 3

Reduction Energies for HeH⁺ with a Hartree-Fock Reference Function
(R = 1.455)

M	$\epsilon^{(m)} = \epsilon_0 + 1/\epsilon^{(m)}[M, M-1]$	% E _{corr} ^f	Brillouin-Wigner Sigma Elements
0	-2.933072	0.00	$\sigma_0 = 0.260175$
1	-2.956818	59.21	$\sigma_1 = 1.723745$
2	-2.966665	83.77	$\sigma_2 = 3.01 \times 10^1$
3	-2.969824	91.64	$\sigma_3 = 6.58 \times 10^2$
4	-2.971765	96.48	$\sigma_4 = 1.64 \times 10^4$
5	-2.972594	98.55	$\sigma_5 = 4.39 \times 10^5$
6	-2.972941	99.41	:
7	-2.973046	99.68	$\sigma_{23} = 5.60 \times 10^{30}$
8	-2.973105	99.82	
9	-2.973138	99.90	
10	-2.973153	99.94	
11	-2.973162	99.97	
12	-2.973168	99.98	
81 (Full CI)	-2.973176		

To obtain a more balanced basis set for potential curve calculations for HeH^+ , basis II was obtained by excluding the π and δ functions, as well as the $4f_0(\text{He})$ and $2s'(\text{H})$ functions, which are least important to the HF solution. These were replaced by $2p_{\pm 1}\text{He}(3.00)$, $2p_{\pm 1}\text{H}(2.00)$, $3d_{\pm 2}\text{He}(3.00)$, and $3d_{\pm 2}\text{H}(2.00)$. This basis results in 61 configurations of ${}^1\Sigma^+$ symmetry. The HF energy at $R = 1.460$ a.u. is -2.93302 a.u. In Figure 1, the HF potential curve and the full CI curve for the $X{}^1\Sigma^+$ ground state can be compared. Since He is a closed shell atom, this is one of the few special cases when the HF potential separates properly, giving a HF atom and a proton. The main feature to be observed is the relative constancy of the correlation correction. In Table 4 the reduction energies through the eighth order are given as a function of R . From the percentages of the correlation energy listed, the convergence with only nine functions is seen to be quite satisfactory, being poorest at $R = 2.50$ a.u. and excellent at large internuclear separations. The first order solution accounts for better than seventy per cent for larger R values. Figure 2 illustrates this as well as the additional improvement given by the second order solution. It should be noted also that the correlation energy is not actually constant but shows a rise at $R = 1.52$ a.u. and decreases slightly as R increases.

The question of the shapes of the potential curves for the reduction energies may also be of importance if one is interested in the vibrational spectra. Thus, to more adequately assess the differences between the curves in Figure 2, a Dunham analysis was applied to each of these approximations using a program written by Beebe [64]. These results are presented in Table 5, along with some comparison values obtained in other calculations by Peyerimhoff [63], Anex [65], Michels [66], and Wolniewicz [67]. Since it is not necessarily obvious just how to treat the separated

TABLE 4
 Reduction Energies for the $X^1\Sigma^+$ State of HeH^+
 as a Function of Internuclear Separation^a
 (Single partitioning with a HF reference function)

$\varepsilon^{(m)}$	R	1.00	1.40	1.46	1.52
ε_0		2.86018	2.93249	2.93302	2.93240
$\varepsilon^{(1)}$		2.88522	2.95771	2.95820	2.95753
$\varepsilon^{(2)}$		2.89392	2.96703	2.96757	2.96693
$\varepsilon^{(3)}$		2.89705	2.97071	2.97130	2.97070
$\varepsilon^{(4)}$		2.89789	2.97179	2.97243	2.97188
$\varepsilon^{(5)}$		2.89807	2.97205	2.97270	2.97217
$\varepsilon^{(6)}$		2.89814	2.97218	2.97283	2.97231
$\varepsilon^{(7)}$		2.89817	2.97223	2.97289	2.97237
$\varepsilon^{(8)}$		2.89818	2.97226	2.97292	2.97240
E_o (Full CI)		2.89818	2.97228	2.97295	2.97243
ΔE		0.03800	0.03979	0.03993	0.04003
% $\Delta E(\varepsilon^{(1)})$		65.89	63.38	63.06	62.78
% $\Delta E(\varepsilon^{(8)})$		100.00	99.95	99.92	99.93

^aAll energies are negative.

TABLE 4 (continued)

$\epsilon^{(m)}$	R	2.00	2.50	3.50	5.00
ϵ_0		2.91057	2.88776	2.86746	2.86218
$\epsilon^{(1)}$		2.93533	2.91243	2.89242	2.88739
$\epsilon^{(2)}$		2.94456	2.92112	2.90039	2.89507
$\epsilon^{(3)}$		2.94835	2.92454	2.90304	2.89752
$\epsilon^{(4)}$		2.94964	2.92557	2.90371	2.89808
$\epsilon^{(5)}$		2.94994	2.92578	2.90387	2.89819
$\epsilon^{(6)}$		2.95011	2.92592	2.90393	2.89822
$\epsilon^{(7)}$		2.95021	2.92600	2.90395	2.89822
$\epsilon^{(8)}$		2.95026	2.92604	2.90396	2.89823
E_0 (Full CI)		2.95033	2.92612	2.90398	2.89823
ΔE		0.03976	0.03836	0.03652	0.03605
$\% \Delta E (\epsilon^{(1)})$		62.27	64.31	68.35	69.93
$\% \Delta E (\epsilon^{(8)})$		99.82	99.79	99.95	100.00

TABLE 4 (continued)

$\varepsilon^{(m)}$	R	6.00	8.00	10.00
ε_0		2.86169	2.86147	2.86142
$\varepsilon^{(1)}$		2.88694	2.88673	2.88668
$\varepsilon^{(2)}$		2.89458	2.89435	2.89430
$\varepsilon^{(3)}$		2.89702	2.89680	2.89678
$\varepsilon^{(4)}$		2.89756	2.89733	2.89727
$\varepsilon^{(5)}$		2.89767	2.89743	2.89738
$\varepsilon^{(6)}$		2.89769	2.89746	2.89740
$\varepsilon^{(7)}$		2.89769	2.89746	2.89741
$\varepsilon^{(8)}$		2.89770	2.89746	2.89741
E_o (Full CI)		2.89770	2.89746	2.89741
ΔE		0.03601	0.03599	0.03599
$\% \Delta E (\varepsilon^{(1)})$		70.12	70.19	70.19
$\% \Delta E (\varepsilon^{(8)})$		100.00	100.00	100.00

FIGURE 1
A Comparison of the HF and Full CI Potential Curves
for the Ground ($X^1\Sigma^+$) State of HeH^+
(Energy vs. Internuclear Separation)
(a.) HF energy
(b.) Full CI energy

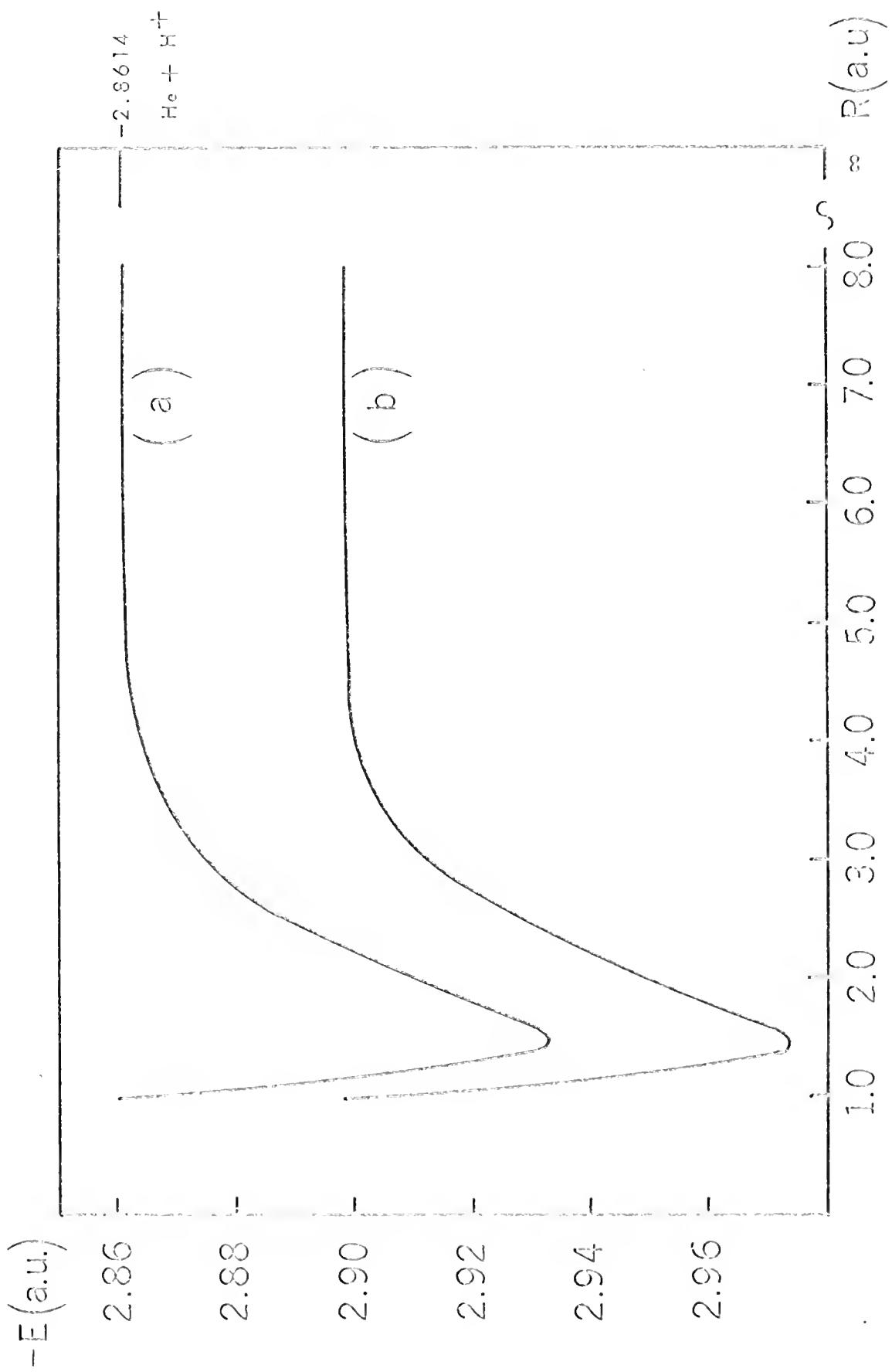
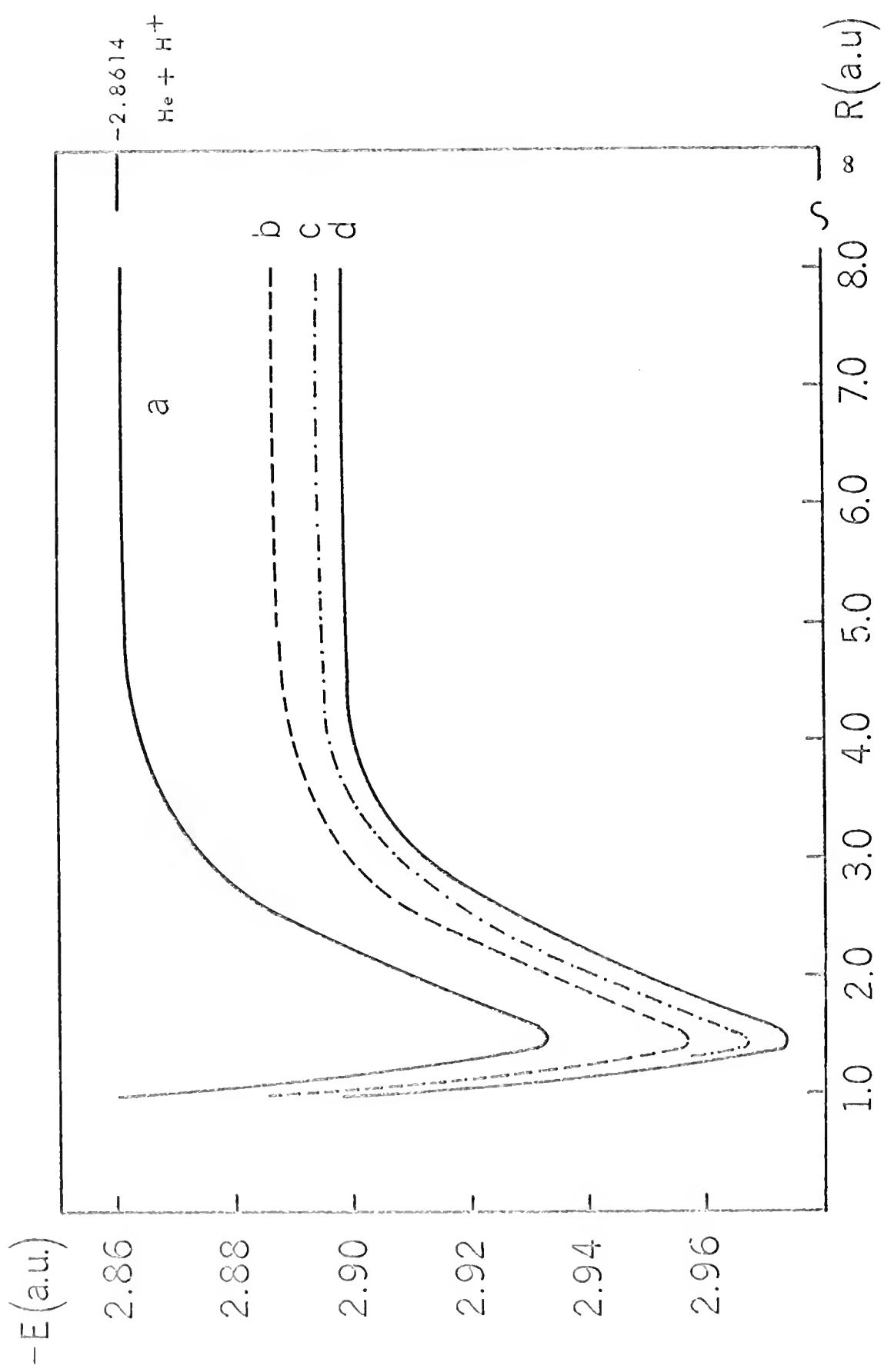


FIGURE 2
Low Order Reduction Energies for the $X^1\Sigma^+$ State of HeH^+
(Energy vs. Internuclear Separation)

- (a.) Unperturbed HF solution
- (b.) First order reduction energy
- (c.) Second order reduction energy
- (d.) Full CI energy



atom limit for the $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$ curves, the values calculated at 10 a.u. were assumed to be a sufficiently close estimate to use in the determination of the dissociation energies. Anex and Wolniewicz used the exact value for the He atom, $E = -2.90372$ a.u., while in Michels' study, the calculated value of $E = -2.87574$ a.u. was the reference point. The values for 10 a.u. should be very close to the calculated atom limit, but even though our dissociation energies seem to be somewhat greater than the comparison values, if any error is present, the magnitude should be slightly underestimated. This is not unexpected, however, since binding energies defined in this manner are not governed by the variational theorem. The lowest total energy found is that of Wolniewicz where $E = -2.97867$ at the minimum given in Table 5. Anex's result is $E = -2.97424$, while Michels reports a value of $E = -2.94373$ a.u.

From the previous spectroscopic parameters in Table 5, it is indicated that despite the dramatic energy decrease, the actual shapes of the potential curves for the first and second order reduction energies appear to be somewhat further from the "best" results than is the HF curve. This effect of some energy improvement destroying agreement with other properties is not uncommon in varieties of perturbation and CI studies. In addition to the shape of the potential curves, another manifestation of this lies in the accuracy of the density distribution which is usually apparent from predictions of the dipole moment or other one-electron properties.

Although the HF wavefunction is known to have certain stability factors [8], for more general cases these questions have not been adequately pursued. However, one can typically expect better agreement for properties more sensitive than the energy when additional flexibility is built into the trial wavefunction. Since these low order reduction solutions have

TABLE 5

Spectroscopic Parameters for Different Approximations to the $X^1\Sigma^+$ State of HeH⁺
(Single partitioning with HF reference function)

	D ₀ (e.v.)	D ₀ (e.v.)	r _e (a.u.)	ω_e (cm ⁻¹)	ω_{ex} (cm ⁻¹)	B _e (cm ⁻¹)	α_e (cm ⁻¹)	K _e (10^{-5} dyynes/cm)
These Calculations								
Hartree-Fock	1.9482	1.7240	1.4576	3615.8	208.0	42.509	3.6344	5.1351
$\varrho^{(1)}$	1.9461	1.7125	1.4553	3767.1	226.0	42.640	3.7714	5.5739
$\varrho^{(2)}$	1.9936	1.7532	1.4574	3877.5	233.8	42.518	3.7617	5.9051
Full CI	2.0555	1.8325	1.4640	3597.1	195.1	42.138	3.4112	5.0820
Comparison Calculations								
HF (a)	1.9437	1.742	1.4553	3257	167	35.30	2.71	5.03
CI (b)	1.931	1.727	1.446	---	---	---	---	---
CI (c)	1.850	---	1.444	3379	313.7	35.85	3.785	5.413
CI (d)	2.0401	---	1.4632	---	---	---	---	---

a Reference [63].

b Reference [65].

c Reference [66].

d Reference [67].

been obtained with only one and two variable parameters, the variational principle apparently overly prejudices the calculation toward an energy reduction with a consequent sacrifice in the accuracy of the potential curves. To avoid these problems, somewhat higher orders must be included. As these are incorporated, the near equality of the eighth order solution and the full CI necessitates that the spectroscopic parameters be essentially identical.

In all of the examples presented in this section, a dramatic energy decrease was given by the low order solutions and essential agreement in the higher orders. In each case, the reduction objective was met with 8 to 12 functions, even though the number of configurations varied from 34 to 82. From the nature of the procedure, it is expected that a similar number of functions should be adequate even with many more configurations. Since we can have only single and double excitations in the systems studied, and the first order solution relative to a HF reference function depends only on the latter, it may be argued that the energy reduction obtained from the geometric formula may not be as pronounced for larger molecules. However, since it is generally conceded that from an energy criterion the double excitations are the most significant, we still anticipate a substantial improvement over the HFR result even for other molecules. To obtain reliable spectroscopic properties, though, it is important to obtain slightly higher order solutions. This feature should also apply to observables that are determined by averaging over the electron density.

VI.3. Results and Discussion: The Selective Reduction Procedure for Multiple States

In order to examine the selective reduction treatment for excited states, the previous 61 configuration, basis II calculation was partitioned to give a ten-function reference space composed of the HF solution and the nine possible singly excited configurations. By solving a CI problem with these ten functions, a set of unperturbed solutions, $|\chi\rangle$, were obtained. From these reference functions, the selective process of (V.2) was used to calculate reduction energies for the ground state and the four lowest excited states of $^1\Sigma^+$ symmetry.

The selectivity of the reduction technique is demonstrated in Table 6. Consistent with the steepest descent argument, the percentages of $\Delta E(i)$ indicate that for each choice of unperturbed solution only the corresponding root is substantially affected. The most efficient reduction is found in the ground state, although the reduction process for the lowest three excited states is also effective. In each of these cases, the first order solution is still seen to be significant.

The ground state perturbation solution obtained with the multi-dimensional reference space is interesting, since, by Brillouin's theorem, the single excitations cannot mix with the HFR ground state, and leave one with exactly the same unperturbed energy as in the single partitioning case. Since more functions are explicitly employed in the multi-dimensional approach, the increased flexibility must of necessity yield some improvement over the single partitioning case for each order. However, the energy decrease is found to be negligibly limited to the fifth decimal place, thus the results in Tables 4 and 5 are the same for all practical purposes. This consequence attests to the fact that the extent of interaction among the elements of the reference space and the contribution from the

TABLE 6

Illustration of Selective Reduction for Multiple States of HeH⁺^a
(Ten-function partitioning; HF and single excitations; R = 1.460)

i	0	1	2	3	4	
Unperturbed Energy (ϵ_i^0)	2.93302	1.91869	1.41766	1.10141	0.61147	
Reference Function						
χ_0^0	$\epsilon^{(1)}$	2.95830	1.91938	1.41900	1.10398	0.61191
	%ΔE(i)	63.31	00.83	02.13	04.79	00.27
	$\epsilon^{(8)}$	2.97294	1.96910	1.41954	1.12804	0.69951
	%ΔE(i)	99.97	60.95	02.99	49.62	53.52
χ_1^0	$\epsilon^{(1)}$	2.93445	1.95689	1.41845	1.10753	0.61159
	%ΔE(i)	03.58	46.19	01.26	11.40	00.07
	$\epsilon^{(8)}$	2.95028	2.00131	1.42300	1.11040	0.73366
	%ΔE(i)	43.22	99.89	08.50	16.75	74.28
χ_2^0	$\epsilon^{(1)}$	2.93651	1.91977	1.44910	1.10166	0.61164
	%ΔE(i)	08.74	01.31	50.06	00.47	00.10
	$\epsilon^{(8)}$	2.93887	1.92574	1.48031	1.11245	0.62182
	%ΔE(i)	14.65	08.52	99.76	20.57	06.29
χ_3^0	$\epsilon^{(1)}$	2.93630	1.92222	1.41778	1.12206	0.61147
	%ΔE(i)	08.21	04.24	00.19	38.48	00.00
	$\epsilon^{(8)}$	2.95176	1.93000	1.43357	1.15461	0.62356
	%ΔE(i)	46.93	13.67	25.33	99.12	07.35

^aAll energies are negative.

TABLE 6 (continued)

i	0	1	2	3	4
χ_4^o	$\varepsilon^{(1)}$	2.93392	1.91880	1.41779	1.10141
	%ΔE(i)	02.25	00.13	00.21	00.00
	$\varepsilon^{(8)}$	2.94321	1.96909	1.42581	1.11437
	%ΔE(i)	25.52	60.94	12.98	24.15
Full CI Solution	(E _i)	2.97295	2.00140	1.48046	1.15508
					0.77597

orthogonal complement is insignificant for this state compared to the mixing of the ground state unperturbed solution, i.e., the HFR result, and the perturbation functions derived from it.

Since basis II is very prejudiced toward the ground state, of the four excited states, only the first, the $A^1\Sigma^+$ state, behaves correctly at large internuclear separation, where it converges reasonably well to a $\text{He}^+(1s) + \text{H}(1s)$ separated atom limit.* The relationship between the $X^1\Sigma^+$ and $A^1\Sigma^+$ curves is illustrated in Figure 3. Although it is not apparent in the figure, from the full CI results listed in Table 7, the first excited state actually shows a slight minimum in the vicinity of $R = 6.5$ a.u. This feature has been previously observed by Michels [66] who more correctly positions the minimum at 5.65 a.u.

In Figure 4 the first and second order reduction energies are compared with the unperturbed result and the full CI solution. As before, the extent of reduction given by the low order solutions is very satisfactory. In fact, $\varepsilon^{(2)}(R)$ apparently converges to the same separated atom limit as the full CI, although $\varepsilon^{(1)}(R)$ does not seem to be able to do so. The eighth order result is observed to be essentially identical to the full solution.

As has been mentioned, the approximate quantities defined in (5.32) are obtained as an intermediate in the computational procedure. In order for these conceptually appealing results to yield reasonable estimates, the interactions arising from the remaining p unperturbed solutions need to be small as was indicated from the ground state discussion. If so, neglecting the mixing should not appreciably affect the spectrum of the results. That this is actually the case in this example is illustrated in

*A correlation diagram for the lower states of HeH^+ has been given by Michels [66].

TABLE 7

Reduction Energies for the $A^1\Sigma^+$ State of HeH^+ as a Function of Internuclear Separation
 (Ten function partitioning; HF and single excitations; Basis III)

$\epsilon_1^{(m)}$	R	2.50	3.50	5.00	6.00	8.00	10.00
ϵ_1^0		2.27919	2.40348	2.44152	2.44409	2.44444	2.44440
$\epsilon_1^{(1)}$		2.32854	2.44445	2.47694	2.47903	2.47922	2.47916
%ΔE	47.95	55.77	62.11	62.92	63.08	63.08	
$\epsilon_1^{(2)}$		2.35976	2.46734	2.49692	2.49927	2.49957	2.49951
%ΔE	78.28	86.93	97.14	99.37	99.98	100.00	
$\epsilon_1^{(8)}$		2.38208	2.47693	2.49855	2.49962	2.49958	2.49951
%ΔE	99.88	99.99	100.00	100.00	100.00	100.00	
E_1 (Full CI)	2.38212	2.47694	2.49855	2.49962	2.49958	2.49951	

FIGURE 3
Full CI Results for the Ground and First Excited States of HeH^+
(Energy vs. Internuclear Separation)

- (a.) Ground ($X^1\Sigma^+$) state
- (b.) Excited ($A^1\Sigma^+$) state

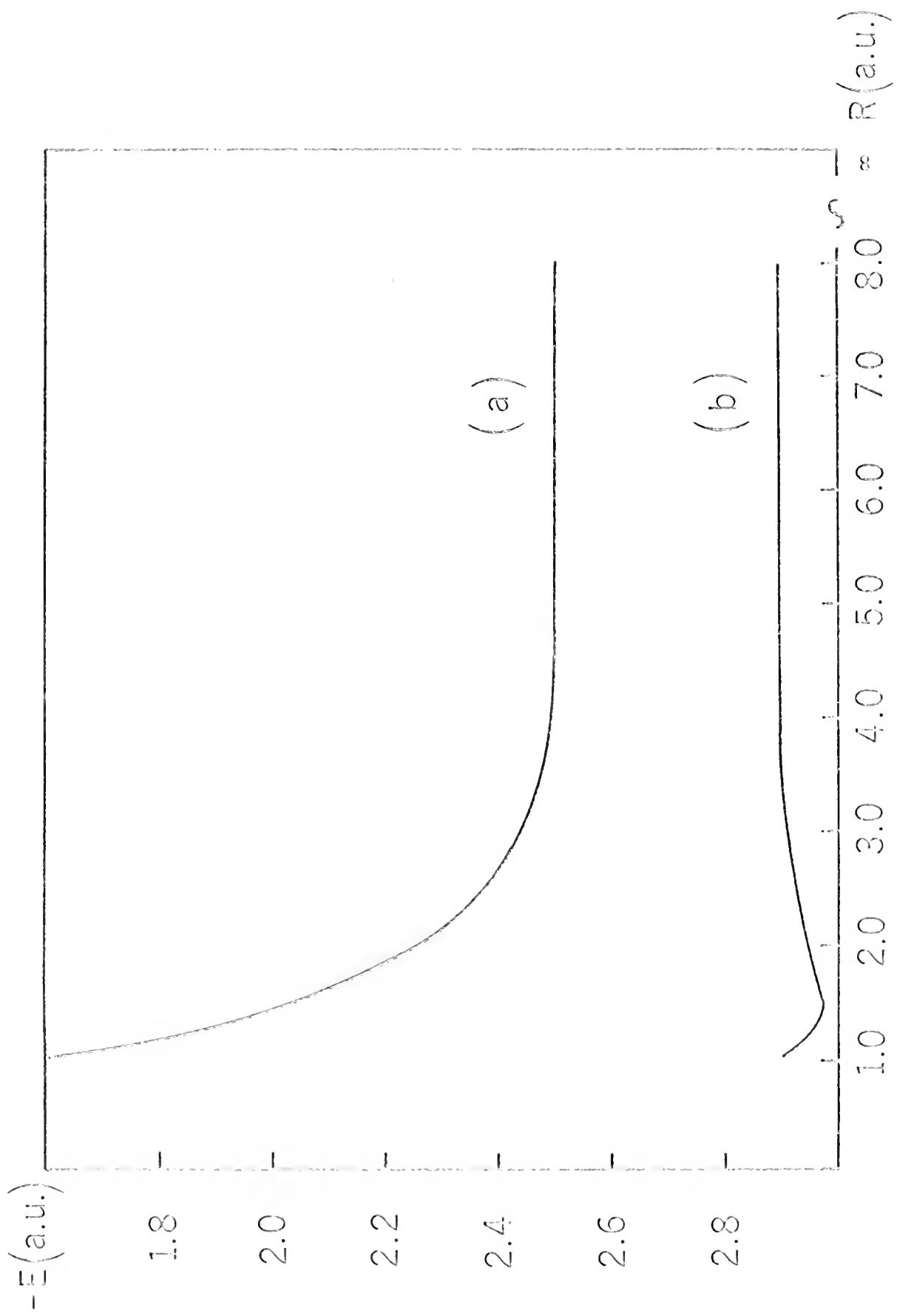


FIGURE 4
Low Order Reduction Energies for the $A^1\Sigma^+$ State of HeH^+
(Energy vs. Internuclear Separation)

- (a.) Ten function CI unperturbed energy
- (b.) First order reduction energy
- (c.) Second order reduction energy
- (d.) Full CI energy

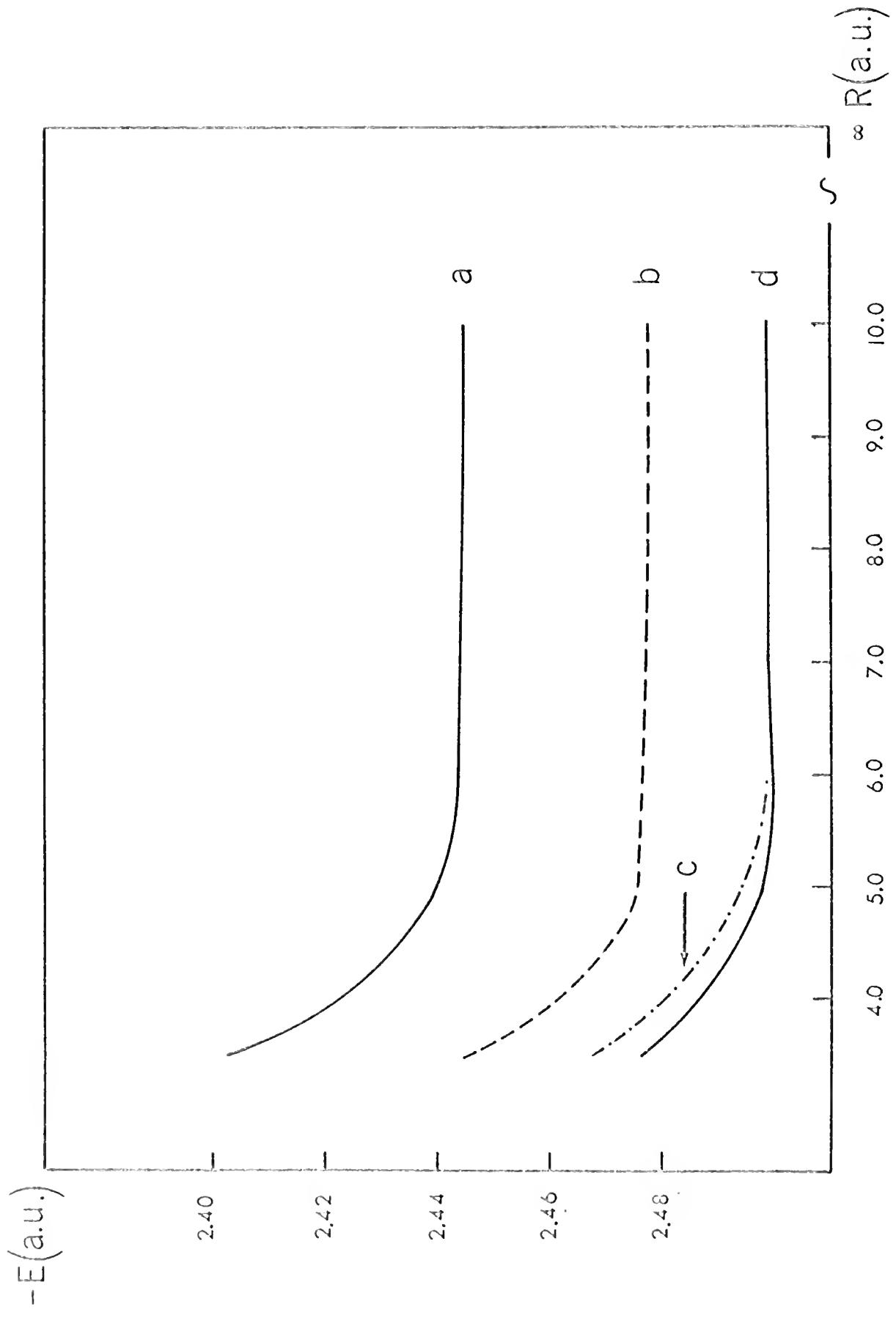


Table 8, where the approximate energies, $\{\tilde{\varepsilon}_i^{(m)}\}$, and the rigorous selectively reduced energies, $\{\varepsilon_i^{(m)}\}$, are compared. Besides the ground state, there is no bounding property for these approximate energies, but as is observed from Table 8, the removal of the remaining reference functions from the space seems to have constrained the problem in such a way that an effect very similar to that perpetrated by the variational principle is present. The resultant eigenvectors also support the validity of this approximation.

An attempt was made to obtain more realistic excited states for HeH^+ that tend to separate correctly at infinite R. At a sacrifice to the ground state, a third basis (III) was assumed that is primarily composed of "best" atom STO's. The basis is $1s_{\text{He}}(2.00)$, $2s_{\text{He}}(0.75)$, $2p_{\text{o He}}(0.75)$, $1s_{\text{H}}(1.00)$, $2s_{\text{H}}(0.50)$, $2p_{\text{o H}}(0.50)$, $1s'_{\text{He}}(3.50)$, $3d_{\text{o He}}(0.60)$, $3d_{\text{o H}}(0.333)$, $2p_{\pm 1 \text{ He}}(0.75)$, $2p'_{\pm 1 \text{ He}}(1.50)$, $2p_{\pm 1 \text{ H}}(0.50)$, $2p'_{\pm 1 \text{ H}}(1.00)$. Since this set contains functions which are able to well approximate the pertinent excited state separated atom limits, $\text{He}^+(1s)$, $\text{He}(1s2s)$, $\text{H}(1s)$, $\text{H}(2s)$, and $\text{H}(2p_{\text{o}})$ [66], the 65 configuration full CI results at 12 a.u. listed in Table 9 show adequate convergence to the experimental separated atom values, -2.500, -2.1460, -2.1250, and -2.1250 a.u., respectively. These curves are illustrated in Figure 5. The change of scale makes the minimum in the $A^1\Sigma^+$ curve apparent, but the relatively broad and deep minimum in the $B^1\Sigma^+$ curve is not readily observed due to R being limited to 8.00 a.u. From the values in Table 9 for larger R, though, this is seen to be the case. The other two states are repulsive.

By comparing these calculations with Michels' values [66], the depths of the wells in the two bound excited states are still somewhat insufficiently described with basis III, though these results are adequate

TABLE 8

Comparison of Rigorous and Approximate Multiple State Reduction Energies for HeH⁺ as a Function of Internuclear Separation
 (Ten function partitioning; HF and single excitations; Basis II)

STATE	R	1.00	1.46	2.00	5.00	10.00
0	$\tilde{\varepsilon}_0^{(1)}$	2.88522	2.95820	2.93533	2.88739	2.88568
	$\varepsilon_0^{(1)}$	2.88529	2.95830	2.93545	2.88750	2.88679
	$\tilde{\varepsilon}_0^{(8)}$	2.89813	2.97278	2.95006	2.89820	2.89738
	$\varepsilon_0^{(8)}$	2.89818	2.97294	2.95032	2.89823	2.89741
1	$\tilde{\varepsilon}_1^{(1)}$	1.56218	1.95664	2.19779	2.47696	2.47916
	$\varepsilon_1^{(1)}$	1.56233	1.95689	2.19793	2.47694	2.47916
	$\tilde{\varepsilon}_1^{(8)}$	1.58897	2.00105	2.25627	2.49855	2.49951
	$\varepsilon_1^{(8)}$	1.58920	2.00131	2.25621	2.49855	2.49951
2	$\tilde{\varepsilon}_2^{(1)}$	1.26363	1.44918	1.46924	1.56815	1.50428
	$\varepsilon_2^{(1)}$	1.26356	1.44909	1.46916	1.56812	1.50420
	$\tilde{\varepsilon}_2^{(8)}$	1.28573	1.47967	1.50847	1.60296	1.53536
	$\varepsilon_2^{(8)}$	1.28582	1.48031	1.51022	1.60302	1.53532
3	$\tilde{\varepsilon}_3^{(1)}$	0.74810	1.12218	1.25367	1.38074	1.46250
	$\varepsilon_3^{(1)}$	0.74800	1.12206	1.25356	1.38062	1.46244
	$\tilde{\varepsilon}_3^{(8)}$	0.77500	1.15608	1.29318	1.41107	1.48318
	$\varepsilon_3^{(8)}$	0.77485	1.15461	1.28917	1.41099	1.48317

TABLE 8 (continued)

STATE	R	1.00	1.46	2.00	5.00	10.00
4	$\tilde{\varepsilon}_4^{(1)}$	0.17290	0.64364	0.96327	1.24013	1.27178
	$\varepsilon_4^{(1)}$	0.17289	0.64363	1.04196	1.24002	1.27178
	$\tilde{\varepsilon}_4^{(8)}$	0.23034	0.77602	0.96325	1.26848	1.29212
	$\varepsilon_4^{(8)}$	0.22625	0.73923	1.01443	1.26833	1.29212

FIGURE 5
Improved Excited $^1\Sigma^+$ States for HeH^+
(Energy vs. Internuclear separation)

- (a.) Fourth excited ($E^1\Sigma^+$) state
- (b.) Third excited ($D^1\Sigma^+$) state
- (c.) Second excited ($B^1\Sigma^+$) state
- (d.) First excited ($A^1\Sigma^+$) state

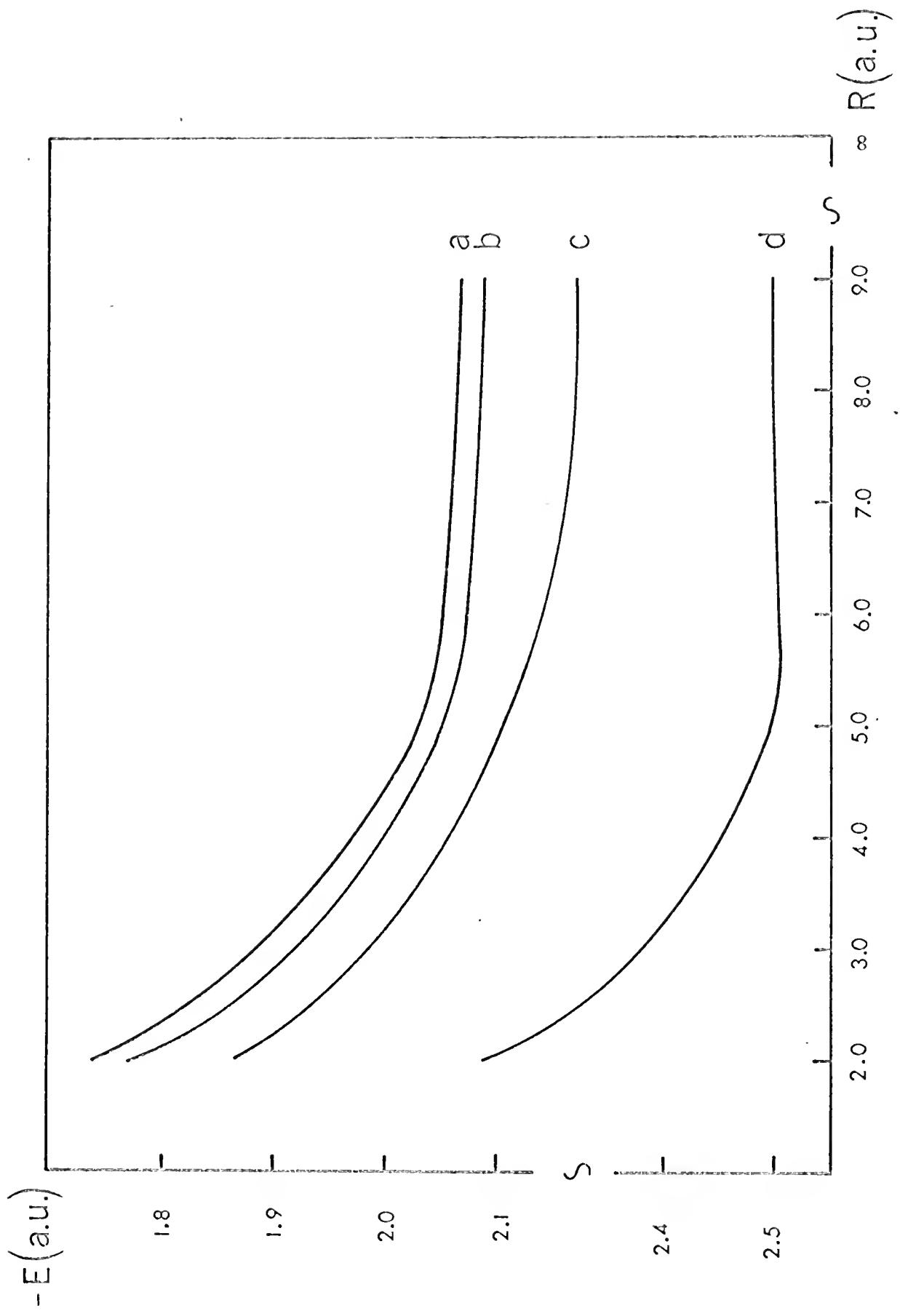


TABLE 9

Reduction Energies for the $^1\Sigma^+$ Excited States of HeH^+ as a Function of Internuclear Separation
(Five function best atom partitioning; Basis III)

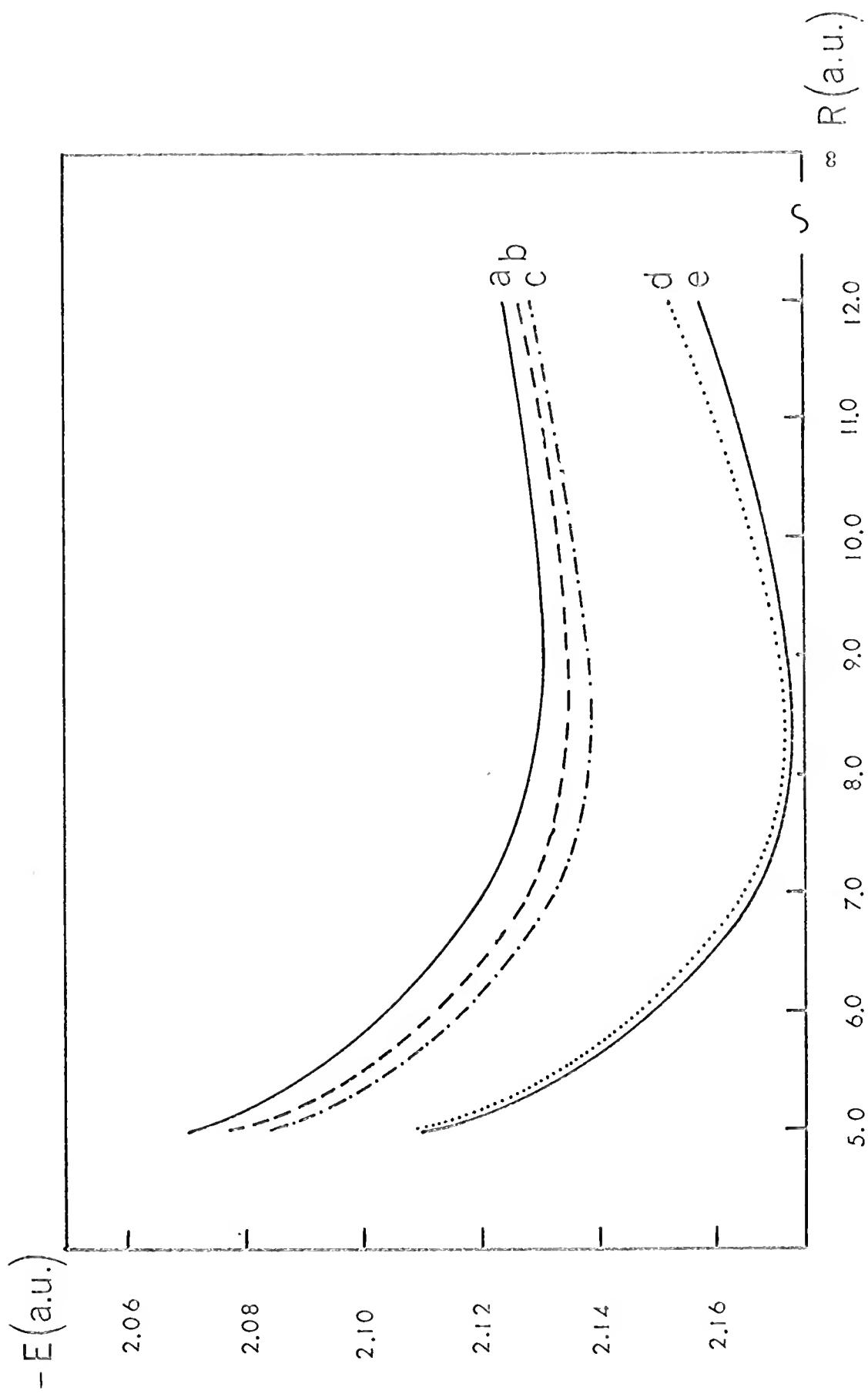
STATE	R	2.00	5.00	6.00	7.00	8.00	9.00	10.00	12.00
\mathcal{E}_1^0		2.20856	2.49823	2.49994	2.50007	2.50004	2.50001	2.50000	2.50000
$\mathcal{E}_1^{(1)}$		2.22888	2.49875	2.50015	2.50019	2.50013	2.50011	2.50010	2.50007
$\% \Delta E(1)$		67.96	32.91	22.83	20.69	24.32	41.67	62.50	87.50
$\mathcal{E}_1^{(8)}$		2.23841	2.49981	2.50086	2.50065	2.50041	2.50025	2.50016	2.50008
$\% \Delta E(1)$		99.83	100.00	100.00	100.00	100.00	100.00	100.00	100.00
\mathcal{E}_1^-		2.23846	2.49981	2.50086	2.50065	2.50041	2.50025	2.50016	2.50008
\mathcal{E}_2^0		1.83900	2.06963	2.10526	2.12269	2.12944	2.13051	2.12888	2.12388
$\mathcal{E}_2^{(1)}$		1.85355	2.07673	2.11253	2.12921	2.13511	2.13534	2.13289	2.12618
$\% \Delta E(2)$		54.74	16.78	16.26	14.58	12.98	11.50	10.08	6.81
$\mathcal{E}_2^{(8)}$		1.86539	2.11137	2.14973	2.16719	2.17292	2.17222	2.16799	2.15541
$\% \Delta E(2)$		99.29	98.63	99.46	99.49	99.56	99.33	98.31	93.42
\mathcal{E}_2^-		1.86558	2.11195	2.14997	2.16742	2.17311	2.17250	2.16866	2.15763

TABLE 9 (continued)

	STATE	R	2.00	5.00	6.00	7.00	8.00	9.00	10.00	12.00
	\mathcal{E}_3^0		1.69788	2.03527	2.04713	2.05341	2.05933	2.06563	2.07171	2.08064
	$\mathcal{E}_3^{(1)}$		1.72344	2.04625	2.05616	2.06272	2.06991	2.07814	2.08367	2.09117
	$\% \Delta E(3)$		37.73	58.06	41.54	36.64	38.08	44.89	44.96	43.53
3	$\mathcal{E}_3^{(8)}$		1.76166	2.05322	2.06868	2.07861	2.08670	2.09274	2.09644	2.10109
	$\% \Delta E(3)$		94.14	94.92	99.13	99.17	98.52	97.27	92.97	84.54
	E_3		1.76563	2.05418	2.06887	2.07882	2.08711	2.09350	2.09831	2.10483
	\mathcal{E}_4^0		1.59658	1.93914	1.96024	1.97429	1.98532	1.99387	1.99998	2.00668
	$\mathcal{E}_4^{(1)}$		1.61902	1.97791	1.99814	2.01216	2.02374	2.03296	2.03952	2.04662
	$\% \Delta E(4)$		16.21	41.71	40.52	42.02	45.54	48.92	50.10	46.02
4	$\mathcal{E}_4^{(8)}$		1.70475	2.02680	2.04710	2.05602	2.06250	2.06587	2.07006	2.07502
	$\% \Delta E(4)$		78.14	94.31	92.87	90.68	91.48	90.10	88.80	78.74
	E_4		1.73501	2.03209	2.05377	2.06442	2.06969	2.07378	2.09890	2.09347

FIGURE 6
Reduction Energies for the $B^1\Sigma^+$ State of HeH^+
(Energy vs. Internuclear Separation)

- (a.) Five function CI unperturbed energy
- (b.) First order reduction energy
- (c.) Second order reduction energy
- (d.) Sixth order reduction energy
- (e.) Full CI energy



estimates. These excited states are very atomic-like and show little correlation. This feature is indicated by the CI eigenvectors which show appreciable mixing for only the first 10 configurations, all of which are obtained from the $1s_{He}$ (2.00) and the nine possible sigma orbitals which may be substituted. Consequently, to improve the results, more extensive experimentation with the basis set is warranted to give CI expansions where the remaining configurations are of more significance.

A partitioning into the first five configurations obtained from the Schmidt orthogonalized spin orbitals that arise from basis III was made to provide an unperturbed reference space for the reduction process. The first five elements of basis III are the spatial orbitals involved in this partitioning, so the unperturbed solution is close to a best atom approximation.

The reduced energies are not as dramatic in first order as previously, which probably signifies the lack of basis adaptability. The eighth order result is still rather good, although it is superior for the lowest $A^1\Sigma^+$ and $B^1\Sigma^+$ states. The first, second, and sixth order solutions for the $B^1\Sigma^+$ state are illustrated in Figure 6. This figure shows better agreement at smaller R, while seeming to approach a somewhat higher separated atom limit than does the full CI result.

This study of the reduced partitioning procedure has in some ways been preliminary, as only two-electron molecules have been considered. Unlike many perturbation approaches, though, the development of this theory always had as its foremost objective improved, or superior, treatments for larger molecules. As such, the formulation is by no means limited to very small systems, but is computationally feasible for more general cases. With the success of this investigation, an extension to more chemically interesting species is justified.

BIBLIOGRAPHY

1. E. Schrödinger, Ann. Physik 79, 361 (1926).
2. M. Born and J. R. Oppenheimer, Ann. Physik 84, 457 (1927).
3. W. Ritz, J. reine angew. Math. 135, 1 (1909).
4. D. R. Hartree, Proc. Cambridge Phi. Soc. 24, 89, 111 (1928); V. Fock, Z. Physik 61, 126 (1930); J. C. Slater, Phys. Rev. 35, 210 (1930).
5. C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951).
6. J. C. Slater, Phys. Rev. 36, 57 (1930).
7. T. A. Koopmans, Physica 1, 104 (1933).
8. C. Moeller and M. S. Plesset, Phys. Rev. 46, 618 (1934); L. Brillouin, Actualites sci. et ind. 71 (1933); J. Goodisman and W. Klemperer, J. Chem. Phys. 38, 721 (1963).
9. P.-O. Löwdin, Phys. Rev. 97, 1474, 1490, 1509 (1955).
10. E. Schmidt, Math. Annln. 63, 433 (1907); A. J. Coleman, Rev. Mod. Phys. 35, 668 (1963).
11. P.-O. Löwdin and H. Shull, J. Chem. Phys. 25, 1035 (1956).
12. E. R. Davidson, J. Chem. Phys. 37, 577 (1962); C. F. Bender and E. R. Davidson, J. Chem. Phys. 47, 360 (1967); 70, 2675 (1966).
13. J. Kouba and Y. Öhrn, J. Chem. Phys. 52, 5387 (1970); 53, 3923 (1970).
14. P.-O. Löwdin, J. Chem. Phys. 19, 1396 (1951).
15. P.-O. Löwdin, J. Mol. Spectry. 10, 12 (1963), PTI; 13, 326 (1964), PTII, PTIII; J. Math. Phys. 3, 969 (1962), PTIV; 3, 1171 (1962), PTV; J. Mol. Spectry. 14, 112 (1964), PTVI; 14, 119 (1964), PTVII; 14, 131 (1964), PTVIII; J. Math. Phys. 6, 1341 (1965), PTIX; Phys. Rev. 139, A357 (1965), PTX; J. Chem. Phys. 43, S175 (1965), PTXI; Perturbation Theory and Its Applications in Quantum Mechanics (Ed. C. H. Wilcox, John Wiley and Sons, New York, 1966), PTXII; Intern. J. Quant. Chem. 2, 867 (1968), PTXIII; P.-O. Löwdin and O. Goscinski, PTXIV, to be published.
16. L. Brillouin, J. Phys. Radium 3, 373 (1932); E. Wigner, Math. naturw. Ang. ungar. Akad. Wiss. 53, 477 (1935).

17. E. Schrödinger, Ann. Physik 80, 437 (1926).
18. E. A. Hylleraas and B. Undheim, Z. Physik 65, 759 (1930).
19. J. K. L. McDonald, Phys. Rev. 43, 830 (1933).
20. N. W. Bazley, Phys. Rev. 120, 144 (1960).
21. N. Aronszajn, "Proceedings of the Symposium on Spectral Theory and Differential Problems," Oklahoma State University, Stillwater, Oklahoma, 1951.
22. E. A. Hylleraas, Z. Physik 65, 209 (1930).
23. P. Lindner and P.-O. Löwdin, Intern. J. Quant. Chem. 2S, 161 (1968).
24. E. Brändas and O. Goscinski, Phys. Rev. 1A, 552 (1970).
25. C. W. Scherr and R. E. Knight, Rev. Mod. Phys. 35, 436 (1963).
26. J. Midtdal, Phys. Math. Univ. Osloensis 32, 1 (1966).
27. H. Padé, Ann. École Super. Normale 9, Suppl. 1 (1892); G. Frobenius, J. reine angew. Math. 90, 1 (1880).
28. G. A. Baker, Jr., Advances in Theoretical Physics (Ed. K. A. Brueckner, Academic Press, New York, 1965).
29. G. A. Baker, Jr., The Padé Approximant in Theoretical Physics (Ed. G. A. Baker, Jr. and J. L. Gammel, Academic Press, New York, 1970). See also the other articles in this volume for various applications of the Padé approximants.
30. H. S. Wall, Analytic Theory of Continued Fractions (Chelsea Publishing Co., New York, 1967).
31. O. Goscinski, Intern. J. Quant. Chem. 1, 769 (1967).
32. J. Nuttall, Phys. Rev. 157, 1312 (1967).
33. O. Goscinski and E. Brändas, Intern. J. Quant. Chem. 5, 131 (1971).
34. B. A. Lippmann and J. Schwinger, Phys. Rev. 79, 469 (1950).
35. E. Brändas and D. Micha, to be published.
36. J. I. Musher, Rev. Mod. Phys. 39, 203 (1967).
37. S. T. Epstein and J. H. Karl, J. Chem. Phys. 44, 4347 (1966).
38. P.-O. Löwdin, Phil. Mag. 5, 46 (1956).
39. B. Riemann, Ger. Math. Werke, 400 (1876); P. Debye, Math. Ann. 67, 535 (1909).

40. R. McWeeny, Proc. Roy. Soc. (London) A235, 496 (1956).
41. A. Dalgarno and A. L. Stewart, Proc. Phys. Soc. (London) 77, 467 (1961); A. Dalgarno and G. W. F. Drake, Chem. Phys. Letters 3, 349 (1969).
42. J. O. Hirschfelder, J. Chem. Phys. 39, 2009 (1962).
43. J. O. Hirschfelder, W. B. Brown, and S. T. Epstein, Advances in Quantum Chemistry, Vol. I (Ed. P.-O. Löwdin, Academic Press, New York, 1964).
44. P. Goldhammer and E. Feenberg, Phys. Rev. 101, 1233 (1956).
45. R. C. Young, L. C. Biedenharn, and E. Feenberg, Phys. Rev. 106, 1151 (1957).
46. H. P. Kelly, Phys. Letters 25A, 6 (1967); Phys. Rev. 144, 39 (1966).
47. J. M. Schulman and J. I. Musher, J. Chem. Phys. 49, 4845 (1968).
48. A. T. Amos, J. Chem. Phys. 52, 609 (1970); A. T. Amos and H. G. Roberts, J. Chem. Phys. 50, 2375 (1969).
49. J. Linderberg and M. Ratner, Chem. Phys. Letters 6, 37 (1970). See also J. Linderberg and Y. Ohn, J. Chem. Phys. 49, 716 (1968); O. Goscinski and B. Lukman, Chem. Phys. Letters 7, 573 (1970).
50. O. Goscinski and E. Brändas, Chem. Phys. Letters 2, 299 (1968); Phys. Rev. 182, 43 (1969).
51. J. H. Choi, "Lower Bounds to Energy Eigenvalues by the Partitioning Technique," Dissertation, University of Florida, 1965; J. H. Choi and D. W. Smith, J. Chem. Phys. 45, 4425 (1966).
52. T. M. Wilson, "Lower Bounds to Eigenvalues of the Schrödinger Equation," Dissertation, University of Florida, 1966; T. M. Wilson, J. Chem. Phys. 47, 3912 (1967); T. M. Wilson and C. E. Reid, J. Chem. Phys. 47, 3920 (1967).
53. J. H. Van Vleck, Electric and Magnetic Susceptibilities (Oxford Press, Oxford, 1932); Phys. Rev. 33, 467 (1929); B. Kirtman, J. Chem. Phys. 49, 3890, 3895 (1968).
54. J. O. Hirschfelder, Intern. J. Quant. Chem. 4S, 257 (1971); P. R. Certain and J. O. Hirschfelder, J. Chem. Phys. 52, 5977 (1970).
55. L. D. Doverspike, R. L. Champion, S. M. Bobbio, and W. G. Rich, Phys. Rev. Letters 25, 909 (1970).
56. W. H. Miller and H. F. Schaefer, J. Chem. Phys. 53, 1421 (1970).
57. J. E. Kouba, "Theoretical Studies of Atom-Atom Interaction," Dissertation, University of Florida, 1970.

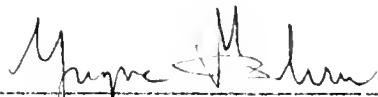
58. C. E. Reid, Preprint No. 196, Quantum Theory Project, University of Florida, Gainesville, Florida; QCPE Program No. 157, Indiana University, Bloomington, Indiana.
59. A. D. McLean, A. Weiss, and M. Yoshimine, Rev. Mod. Phys. 32, 212 (1960).
60. C. A. Coulson, Trans. Faraday Soc. 33, 1479 (1937).
61. R. J. Bartlett and E. J. Brändas, Intern. J. Quant. Chem. 5S, 0000 (1971); E. J. Brändas and R. J. Bartlett, Chem. Phys. Letters 8, 153 (1971).
62. G. L. Bendazzoli, O. Goscinski, and G. Orlandi, Phys. Rev. 2A, 2 (1970).
63. S. Peyerimhoff, J. Chem. Phys. 43, 998 (1965).
64. N. Beebe, Program SPECPR, Quantum Theory Project, University of Florida, Gainesville, Florida.
65. B. G. Anex, J. Chem. Phys. 38, 1651 (1963).
66. H. H. Michels, J. Chem. Phys. 44, 3834 (1966).
67. L. Wolniewicz, J. Chem. Phys. 43, 1087 (1965).

BIOGRAPHICAL SKETCH

Rodney Joseph Bartlett was born March 31, 1944, in Memphis, Tennessee, where he graduated from Whitehaven High School in June, 1962. In May, 1966, he received cum laude the degree of Bachelor of Science from Millsaps College with a double major in chemistry and mathematics. He entered the Graduate School of the University of Florida in September, 1966, as an NDEA Title IV Fellow, a position he held until August, 1969. From September, 1969, until June, 1971, he held an IBM Pre-Doctoral fellowship.

Rodney Joseph Bartlett is married to the former Beverly Jean Featherston and is the father of one son. He is a member of Omicron Delta Kappa and has been named to Who's Who Among Students in American Universities and Colleges.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Yngve Ohrd, Chairman
Professor of Chemistry and Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Per-Olov Lowdin, Co-Chairman
Professor of Chemistry and Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Willis E. Person
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Thomas L. Bailey
Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John P. Shew

Zoran R. Pop-Stojanovic
Associate Professor of Mathematics

This dissertation was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

June, 1971

H. H. St. Hwy

Dean, Graduate School



UNIVERSITY OF FLORIDA



3 1262 08553 7925